3. Simulation Models and Tools

- Broadcast frequency: **10Hz**, which means 100ms. Here **plain EEBL** means without rebroadcast.
- **P-persistence**: The rebroadcast decision is taken with a probability *P* which depends on the distance from the sender. → **EEBLR**
- Every 100ms, the queue is empted and **a single frame** aggregating all messages is sent. → **EEBLA**

3. Simulation Models and Tools

The safe gap depends on the speed of the current vehicle and is computed:

$$s_{safe} = T_{ACC} \cdot v_i + \varepsilon_{ACC}$$

If actual gap is lower than the safe gap, the follower brakes; else computes acceleration

$$a^{ACC} = \frac{v_{i+1}^2 - v_i^2}{2 \cdot (s_{actual} - s_{safe})}$$

4. Network Level Results

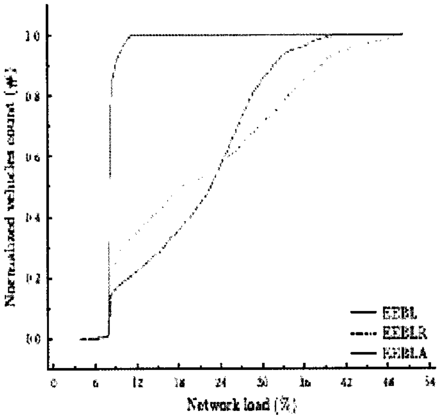


Figure 4: Maximum channel load ρ_i^{max} observed by each vehicle during the simulation for the five-lane scenario for plain EEBL, EEBLR and EEBLA

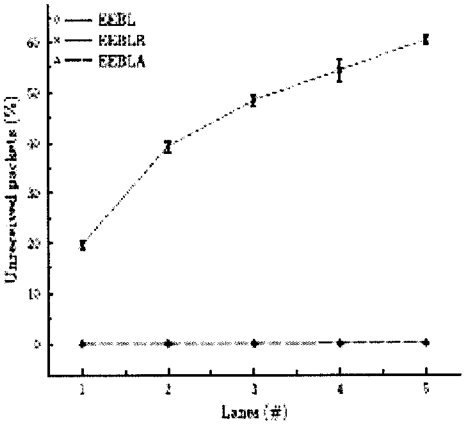


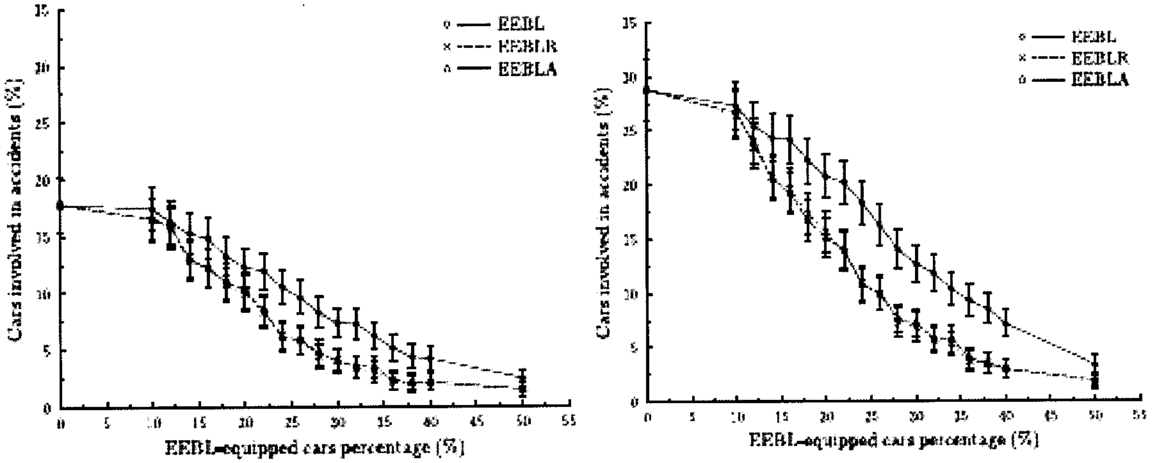
Figure 5: l_{uf} as a function of the number of lanes for plain EEBL, EEBLR and EEBLA

4. Network Level Results

Two metrics to measure the network performance:

- Channel load at each station $\boldsymbol{\rho_i(t)}$**;**
- The percentage of frames $\boldsymbol{l_{uf}}$ that are not received by any station.

5. Application Performance

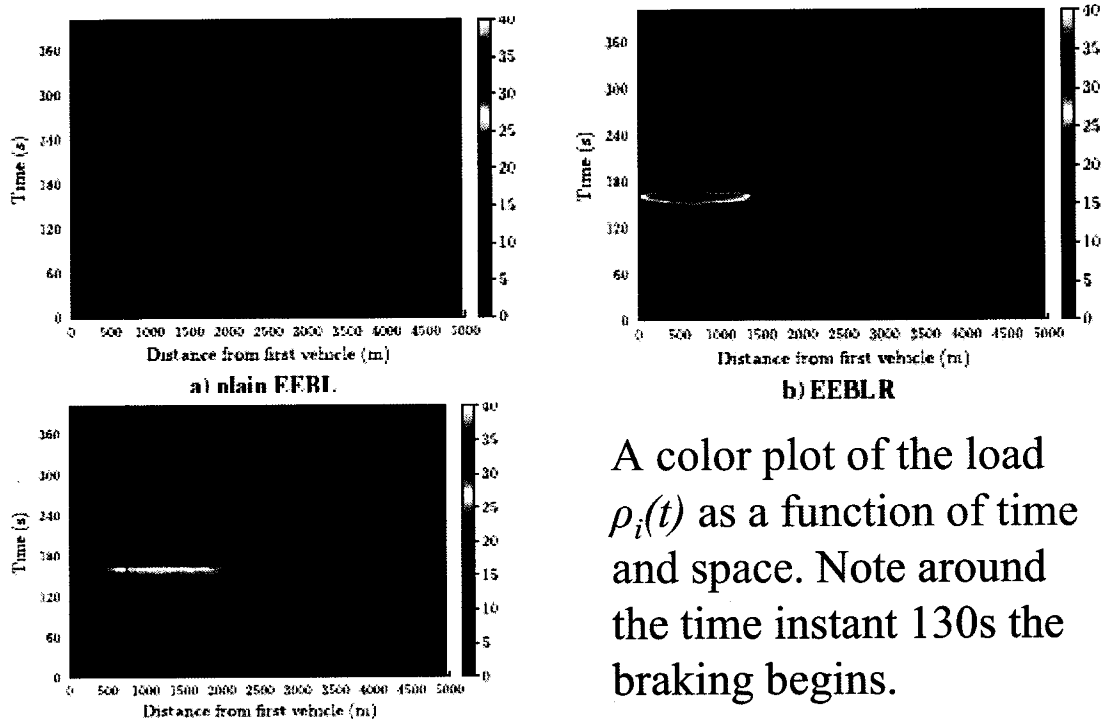


a) 130 km/h initial speed

b) 150 km/h initial speed

Percentage of cars involved in accidents vs. Market Penetration Rate for single-lane tests for different protocols and different average speeds

4. Network Level Results



a) plain EEBL

b) EEBLR

c) EEBLA

A color plot of the load $\rho_i(t)$ as a function of time and space. Note around the time instant 130s the braking begins.

5. Application Performance

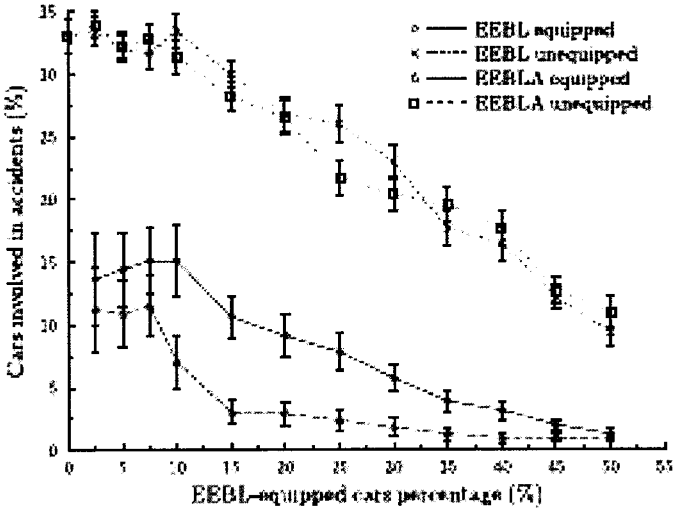
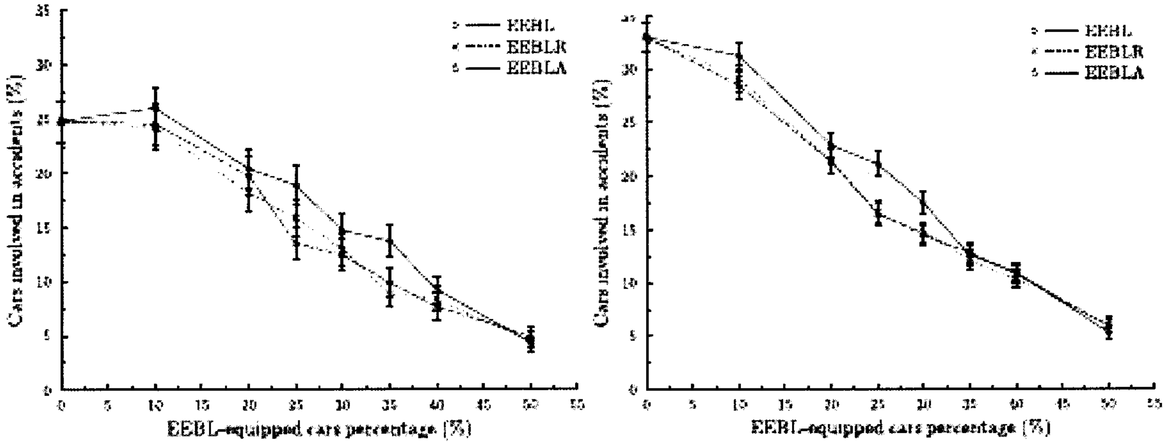


Figure 9: Split-down of equipped an unequipped cars involved in accidents vs. MPR; tests in a five lane-scenario for EEBL and EEBLA, average speed 130 km/h

5. Application Performance



a) two-lane scenario

b) five-lane scenario

Percentage of cars involved in accidents vs. Market Penetration Rate for multi-lane tests, average speed 130 km/h, two and five-lane scenarios

6. Conclusions and Future Work

- **The time-space analysis of the channel** load is novel, as it gives insight in the dynamics of the joint network and application evolution .

Appendix XXV. Lijian Xu et al., *Communication Information Structures and Contents for Enhanced Safety of Highway Vehicle Platoons.*

Communication Information Structures and Contents for Enhanced Safety of Highway Vehicle Platoons

Lijian Xu, *Student Member, IEEE*, Le Yi Wang, *Fellow, IEEE*, George Yin, *Fellow, IEEE*, Hongwei Zhang, *Senior Member, IEEE*

Abstract—**Highway platooning of vehicles has been identified as a promising framework in developing intelligent transportation systems. By autonomous or semi-autonomous vehicle control and inter-vehicle coordination, an appropriately managed platoon can potentially offer enhanced safety, improved highway utility, increased fuel economy, and reduced emission. This paper is focused on quantitative characterization of impact of communication information structures and contents on platoon safety. By comparing different information structures which combine front sensors, rear sensors, and wireless communication channels, and different information contents such as distances, speeds, and drivers' actions, we reveal a number of intrinsic relationships between vehicle coordination and communications in platoons. Typical communication standards and related communication latency are used as benchmark cases in our study. The findings of this paper provide useful guidelines in sensor selections, communication resource allocations, and vehicle coordination.**

Index Terms—**Highway platoons, vehicle safety, communication systems, communication latency, autonomous vehicles.**

I. Introduction

Highway platooning of vehicles has been identified as a promising framework in developing intelligent transportation systems [1], [2]. By autonomous or semi-autonomous vehicle control and inter-vehicle coordination, an appropriately managed platoon can potentially offer enhanced safety, improved highway utility, increased fuel economy, and reduced emission. In a platoon formation and maintenance, high-level distributed supervisors adjust vehicle spatial distributions based on inter-vehicle information such that roadway utilization is maximized while the risk of collision is minimized or avoided. Controllers at vehicle levels, sensors, and communication systems interact intimately in vehicle platoon formation and control. This paper investigates several key issues in such interactions.

Platoon control has drawn substantial attention lately [3], [4]. During the 90s, there were substantial contributions on platoon control, including PATH projects [5], [6], FleeNet, among others. Intelligent platoon control algorithms were introduced with demonstration and experimental validation [7], [8]. The most common objectives in platoon control are safety, string stability, and team coordination [9], [10]. Early studies of platoon control were not communication focused, due to less-advanced communication systems at that time. In our recent work [11], [12], a weighted and constrained consensus control method was introduced to achieve platoon formation and robustness. At present, on-board front radars are used in vehicle distance measurements. [12] employs convergence rates as a performance measure to evaluate benefits of different communication topologies in improving platoon formation, robustness, and safety.

Communication channels insert new dynamic subsystems into control loops. Impact of communication systems on feedback loops can be treated as added uncertainty such as delays and errors [13], [14], [15]. In terms of coordination of control and communication systems in a platoon, some intrinsic questions arise: (1) How much improvement of safety can be achieved by including communication channels? (2) What information should be communicated? What are the values of such information? (3) How will communication uncertainties such as latency, packet loss, and error affect safety?

This paper aims to answer these questions with quantitative characterization. To facilitate this exploration, we consider various information structures: (1) front radars only, (2) combined radars and wireless communications. In addition, we investigate the information contents: (1) distances only, (2) distance and speed, (3) additional early warning of the driver's braking action. Typical communication standards such as IEEE 802.11p and related communication latency are used as benchmark cases in this study. The findings of this paper will be useful to guide design of information infrastructures, information contents, control strategies, and resource allocations in platoon control problems.

The rest of the paper is organized into the following sections. Section II introduces the basic platoon control problem and safety issues. Section III defines control strategies and sets up evaluation scenarios for comparative studies of different information structures and contents. Our studies start with safety analysis in Section IV. Under some simplified scenarios, basic relations are derived, including speed-distance relationship for safe stopping distance and collision avoidance, distance progression in a platoon, and delay-distance functions

Lijian Xu and Le Yi Wang are with the Department of Electrical and Computer Engineering, Wayne State University, Detroit, Michigan 48202. Email: dy0747@wayne.edu (Xu), lywang@wayne.edu (Wang)

George Yin is with the Department of Mathematics, Wayne State University, Detroit, MI 48202. Email: gyin@math.wayne.edu

Hongwei Zhang is with the Department of Computer Science, Wayne State University, Detroit, MI 48202. Email: hongwei@wayne.edu

This research was supported in part by the National Science Foundation under grant CPS-1136007.

for communication latency. Section V details typical communication scenarios. Communication latency characterization and related experimental data are presented. Section VI investigates impact of information structure by comparing radar-based distance sensing and communications. Front radars are the current commercial automotive technology. By expanding information structures to include wireless communication networks, improvement on safety is quantitatively studied. The roles of information contents are explored in Section VII, in which improvements on safety by including more information on vehicle speeds and drivers' actions are studied. Section VIII investigates impact of communication latency and uncertainties on vehicle safety. Typical scenarios of communication latency, radar resolution, and Doppler frequency shifting are considered. Finally, Section IX discusses implications of the results of this paper and points out some potential extensions.

II. Vehicle Dynamics and Platoon Information Structure

This paper is concerned with inter-vehicle distance control in a highway platoon. For clarity of investigation, we use simplified, generic, but representative vehicle dynamic models from [16]

$$m\dot{v} + f(v) = F, \tag{1}$$

where m (Kg) is the consolidated vehicle mass (including vehicle, passengers, etc.), v is the vehicle speed (m/s), $f(v)$ is a positive nonlinear function of v representing resistance force from aerodynamic drag and tire/road rolling frictions, and F (Newton or Kg-m/s^2) is the net driving force (if $F > 0$) or braking force (if $F < 0$) on the vehicle's gravitational center. Typically, $f(v)$ takes a generic form $f(v) = av + bv^2$, where the coefficient $a > 0$ is the tire/road rolling resistance, and $b > 0$ is the aerodynamic drag coefficient. These parameters depend on many factors such as the vehicle weight, exterior profile, tire types and aging, road conditions, wind strength and directions. Consequently, they are determined experimentally and approximately. This paper focuses on longitude vehicle movements within a straight-line lane. Thus, the vehicle movement is simplified into a one-dimensional system.

Vehicles receive platoon movement information by using sensors and communication systems. We assume that radars are either installed at front or rear of the vehicle. The raw data from the radars are distance information between two vehicles. Although it is theoretically possible to derive speed information by signal processing (derivatives of the distances), this paper works with the direct information and leaves signal processing as part of control design. As a result, radar information is limited to distances. In contrast, a communication channel from vehicle i to vehicle j can transmit any information that vehicle i possesses. We consider the following information contents for transmission: (1) vehicle i's distance that is measured by its front sensor, (2) vehicle i's speed, which is available by its own speedometer, (3) vehicle i's braking action. Information structures are depicted in Fig. 1. A vehicle may receive information from its front distance sensor (on its distance to the front vehicle), or its rear sensor (on its distance to the vehicle behind it), or wireless communication channels between two vehicles. The wireless communication channels may carry different information contents such as distance, speed, driver's action, etc.

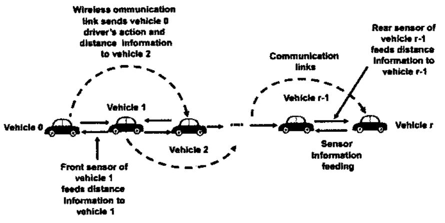

Wireless ommunication link sends vehicle 0 driver's action and distance information to vehicle 2
Communication links
Rear sensor of vehicle r-1 feeds distance information to vehicle r-1
Vehicle 1
Vehicle r-1
Vehicle 0
Vehicle r
Vehicle 2
Sensor information feeding
Front sensor of vehicle 1 feeds distance information to vehicle 1

Fig. 1. Information structures.

For concreteness, we use a basic three-car platoon to present our key results. Although this is a highly simplified platoon, the main issues are revealed clearly in this system. Three information structures are studied, shown in Fig. 2. "Information Structure (a)" employs only front sensors, implying that vehicle 1 follows vehicle 0 by measuring its front distance d_1, and then vehicle 2 follows vehicle 1 by measuring its front distance d_2. For safety consideration, this structure provides a baseline safety metric for comparison with other information structures. "Information Structure (b)" provides both front and rear distances. Then "Information Structure (c)" expands with wireless communication networks.

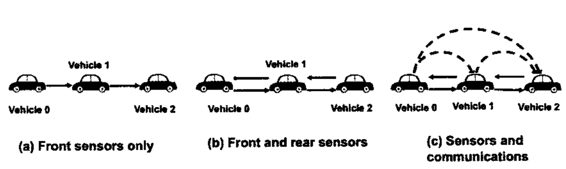


Fig. 2. Three main information structures: (a) Only front distance information is available for vehicle control. (b) Both front and rear distances are available. (c) Additional information is transmitted between vehicles.

Although we employ a three-car platoon for simplicity, it forms a generic base for studying platoon safety issues for more general platoons. This is graphically explained in Fig. 3. Here the vehicles in between the leading vehicle and the vehicle of interest are grouped as one sub-platoon. We treat this sub-platoon as one vehicle and this leads to the generic structure of Fig. 2. This also implies that the communication distance between the two vehicles may be high.

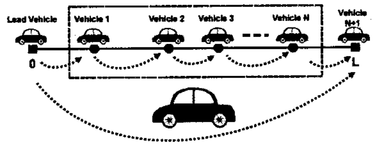


Fig. 3. Grouping vehicles.

The platoon in Fig. 2 has the following local dynamics,

$$\begin{cases} \dot{v}_0 &= \frac{1}{m_0}(F_0 - (a_0 + b_0 v_0^2)) \\ \dot{v}_1 &= \frac{1}{m_1}(F_1 - (a_1 + b_1 v_1^2)) \\ \dot{v}_2 &= \frac{1}{m_2}(F_2 - (a_2 + b_2 v_2^2)) \\ \dot{d}_1 &= v_0 - v_1 \\ \dot{d}_2 &= v_1 - v_2, \end{cases} \qquad (2)$$

where F_0 is the leading vehicle's driving action. F_1 and F_2 are local control variables. Since the vehicle lengths are fixed and can be subtracted from distance calculations, in this formulation a vehicle is considered as a point mass without length.

III. Control and Evaluation Scenarios

A. Feedback Control

For safety consideration, the inter-vehicle distances d_1 and d_2 have a minimum distance $d_{min} > 0$. To ensure that vehicles 1 and 2 have sufficient distances to stop when the leading vehicle 0 brakes, a cruising distance d_{ref} is imposed. Apparently, the larger d_{ref}, the safer the platoon, under any fixed control strategies. However, a larger d_{ref} implies more occupation of the highway space, and less efficiency in highway usage. As a result, it is desirable to use as small d_{ref} as possible without compromising the safety constraint.

There are numerous vehicle control laws which have been proposed or commercially implemented [17], [18]. Since the focus of this paper is on impact of information structures and contents rather than control laws, we impose certain simple and fixed control laws. For safety consideration, we concentrate on the case when the distance is below the nominal value $d < d_{ref}$. The control law involves a normal braking region (small slope) and an enhanced braking region of a sharp nonlinear function towards the maximum braking force, as shown in Fig. 4. We denote this function as $F = g_1(d)$.

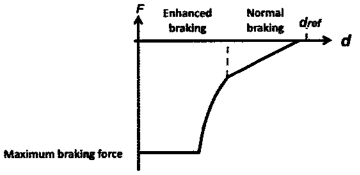


Fig. 4. Braking functions based on distance information.

Similarly, if vehicle i's speed information is transmitted to another vehicle j (behind i), the receiving vehicle can use this information to control its braking force. This happens when $v_j > v_i$. The larger the difference, the stronger the braking force. This control strategy may be represented by a function $F = g_2(v_j - v_i)$, shown in Fig. 5.

B. Evaluation Scenarios

To investigate impact of information structures and contents on platoon safety, we need a reasonable platform to comparative studies. Since vehicle safety involves so many factors, we must define a highly simplified platform in which only key elements are represented. For this reason, we define the following basic scenarios.

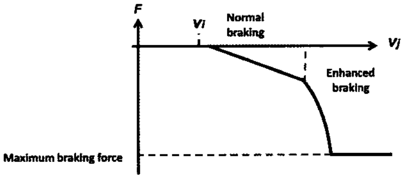


Fig. 5. Braking functions based on speed information.

We use some typical vehicle data from [16]. Under the MKS (metre, kilogram, second) system of units, the vehicle mass m has the range $1400 - 1800$ Kg, the aerodynamic drag coefficient b has the range $0.35 - 0.6$ Kg/m. During braking, a (as the rolling resistance) is changed to tire/road slipping, which is translated into the braking force F (negative value in Newton). As a result, a is omitted.

Three identical cars form a platoon as in Fig. 2. The vehicle masses are $m_0 = m_1 = m_2 = m = 1500$ Kg. The aerodynamic drag coefficients $b_0 = b_1 = b_2 = 0.43$. The nominal inter-vehicle distance $d_{ref} = 40$ m. The cruising platoon speed is 25 m/s (about 56 mph). The road condition is dry and the maximum braking force is 10000 N. This implies that when the maximum braking is applied (100% slip), the vehicle will come to a stop in 3.75 second. The braking resistance can be controlled by applying controllable forces on the brake pads.

The feedback control function $F = g_1(d)$ is depicted in Fig. 6. The actual function is

$$\max\{k_1(d - dref) + k_2(d - dref)^3, -F_{max}\} \qquad (3)$$

where $d_{ref} = 40$ (m), $k_1 = 50$, $k_2 = 4$, $F_{max} = 10000$ (N). The function applies smaller braking force when the distance is only slightly below the reference value, but increases the braking force more dramatically in a nonlinear function when the distance reduces further until it reaches the maximum braking force. We comment that if one views the braking function purely from safety aspects, it is desirable to impose the maximum braking as soon as the distance drops. This, however, will compromise drivability and smoothness of platoon operation. In fact, the braking function of Fig. 6 is already on the aggressive side.

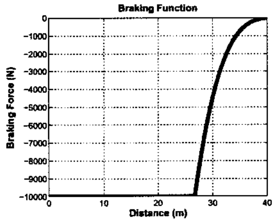


Fig. 6. Braking function for Example 4.

To see this, consider the slow braking condition: Suppose that the leading vehicle applies a braking force 1000 N, which brings it to a stop from 25 m/s in 37.5 second. The distance trajectories of d_1 and d_2 are shown in Fig. 7. In this case, the minimum distances are 30.9 m for d_1 and 24.2 m for d_2. This is acceptable for safety. On the other hand, the transient period shows oscillation, indicating that the braking action has been aggressive already under normal driving conditions.

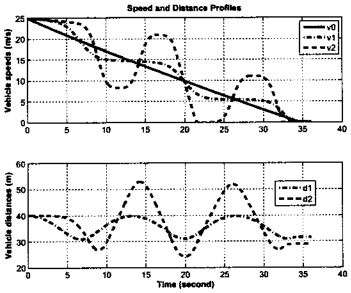


Fig. 7. Distance trajectories under slow braking.

For evaluations, we will use the fast braking scenario defined as follows.

Fast Braking: The leading vehicle uses a braking force 5000 N. If the cruising speed of the platoon is 25 m/s, then this braking force brings the leading vehicle to a stop from 25 m/s in 7.5 second.

In some derivations, we also use the extreme case in which the maximum braking force 10000 (N) is applied. This is for the worst-case analysis. But the Fast Braking case is representative for understanding safety issues. In this paper, the minimum vehicle distance $d_{min} = 15$ (m) is used to distinguish "acceptable" and "unsafe" conditions. When a distance is reduced to 0, a collision occurs.

IV. Safety Analysis

We conduct safety analysis under the scenario specified in Section III-B. Some simplifications will be made so that explicit expressions can be derived to clarify the main underlying safety issues.

We observe that under this braking force, the influence of the tire/road resistance and aerodynamic drag force bv^2 is relatively small. a is proportional to the tire deformation and inversely proportinal to the radius of the loaded tire. The rolling resistance of a normal car 1500 kg on convrete with rolling coefficient 0.01 can be estimated:

$$F_r = 0.01(1500kg)(g) = 0.03(1500kg)(9.81m/s^2) = 147(N), \tag{4}$$

When $b = 0.43$ and $v = 25$ m/s, the aerodynamic drag force is 268.75 (N). This is only 8.3% of the braking force. In the subsequent development, we omit the aerodynamic drag force in our derivations, but include it in all simulation studies.

Assuming that the platoon cruising speed is $v_0(0) = v_1(0) = v_2(0) = 25$ (m/s) and the leading vehicle brakes at $t = 0$ with $F_0 = -\alpha$, where α is a constant (for the Fast Braking, $\alpha = 5000$ (N); and the worst-case $\alpha = F_{max} = 10000$ (N)). The braking function (3) is used. It follows that the dynamics of the three-car platoon are

$$\left\{ \begin{array}{lcl} \dot{v}_0 & = & -\frac{\alpha}{m} \\ \dot{v}_1 & = & -\frac{g_1(d_1)}{m} \\ \dot{v}_2 & = & -\frac{g_1(d_2)}{m} \\ \dot{d}_1 & = & v_0 - v_1 \\ \dot{d}_2 & = & v_1 - v_2, \end{array} \right. \tag{5}$$

with the initial conditions $v_0(0) = v_1(0) = v_2(0) = 25$ (m/s) and $d_1(0) = d_2(0) = d_{ref} = 40$ (m).

A. Safety Regions

In a platoon, usually vehicle 2 acts later than vehicle 1 due to information cascading structures (vehicle 1 sees the slowdown of the leading vehicle before vehicle 2). Suppose that after vehicle 1 applied the maximum braking force at an earlier time, vehicle 2 starts to apply the maximum braking force at t_0.

Theorem 1: Assume that $v_1(t_0) < v_2(t_0)$. Denote $\eta = v_2^2(t_0) - v_1^2(t_0)$, and $\delta = d_2(t_0)$. The final distance is

$$d_2^{final} = \delta - \frac{\eta m}{2F_{max}}.$$

Proof: For $t \geq t_0$, the two vehicles have the dynamics $\dot{v}_1 = -\frac{F_{max}}{m}$, $\dot{v}_2 = -\frac{F_{max}}{m}$, which implies $v_1(t) = v_1(t_0) - \frac{F_{max}}{m}(t - t_0)$, $v_2(t) = v_2(t_0) - \frac{F_{max}}{m}(t - t_0)$.

Vehicle 1 stops after travelling the total stoping time $v_1(0)m/F_{max}$ and the total length $\Delta_1 = v_1^2(t_0)m/(2F_{max})$. Similarly, the total length travelled by vehicle 2 to a complete stop is $\Delta_2 = v_2^2(t_0)m/(2F_{max})$. Thus, the final distance is

$$d_2^{final} = \delta - \frac{(v_2^2(t_0) - v_2^1(t_0))m}{2F_{max}} = \delta - \frac{\eta m}{2F_{max}}.$$

□

For any given final distance $d_2^{final} = C$, the function

$$\eta = \frac{2F_{max}}{m}(\delta - C)$$

defines the iso-final-distance line on the $\delta - \eta$ space, shown in Fig. 8, in which the acceptable region and collision avoidance region are also marked.

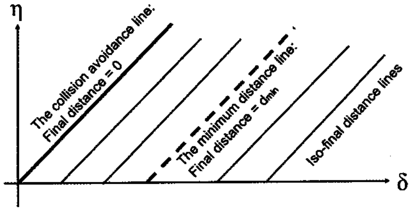


Fig. 8. $\delta - \eta_v$ lines under a given final distance. Acceptable safety regions and collision avoidance regions can be derived from such curves.

B. Platoon Distance Progression

A platoon consists of many vehicles. Under typical information structures, there is a phenomenon of inter-vehicle distance progression that must be considered in platoon management.

Assumption 1: (1) At $t = 0$, the platoon of n following vehicles is at the cruising condition with equal distance d_{ref} and speed $v(0)$. (2) The information on the braking action $F_0 = -F_{max}$ of the leading vehicle at $t = 0$ is passed to the following vehicles in a progressive manner: For $t > 0$, $F_1(t) < F_2(t) < \cdots < F_n(t)$, except when the braking forces reach the saturation values -10000 (N), in which case, one may have $F_1(t) = F_2(t)$, etc. (3) Suppose that vehicle j starts to apply the maximum braking force at t_j. We assume that $t_1 < t_2 < \cdots < t_n$.

Theorem 2: Under Assumption 1, the total travel length L_j of vehicle j before a complete stop satisfies

$$L_0 = \frac{v(0)m}{2F_{max}} < L_1 < L_2 < \cdots L_n.$$

The minimum final distance is

$$\min_{j=1,\ldots n} d_j^{final} = d_{ref} - \max_{j=1,\ldots,n} (L_j - L_{j-1}).$$

Proof: The expression $L_0 = \frac{v(0)m}{2F_{max}}$ is proved in Theorem 1.

Let the braking force for vehicle j be $-f_j(t)$ with $f_j > 0$. The speed profile is

$$v_j(t) = v(0) - \int_0^t \frac{f_j(\tau)}{m} d\tau.$$

The total travel time T_j satisfies

$$\int_0^{T_j} \frac{f_j(\tau)}{m} d\tau = v(0).$$

The total length travelled by vehicle j until a complete stop is

$$L_j = \int_0^{T_j} v_j(t)dt = v(0)T_j - \int_0^{T_j} \int_0^t f_j(\tau)d\tau dt.$$

Under Assumption 1, we have the inequalities

$$v_1(t) < v_2(t) < \cdots < v_n(t), t > 0 \qquad (6)$$

which implies that

$$T_1 < T_2 < \cdots T_n. \qquad (7)$$

These imply

$$L_1 < L_2 < \cdots L_n.$$

Now, the final distance d_j^{final} is

$$d_j^{final} = d_{ref} - (L_j - L_{j-1})$$

which implies that

$$\min_{j=1,\ldots,n} d_j^{final} = d_{ref} - \max_{j=1,\ldots,n} (L_j - L_{j-1}).$$

This completes the proof. □

C. Delay-Distance Relationship

This paper concentrates on communication latency and its impact on vehicle safety. In this subsection, a relationship between the communication delay time and its detrimental effect on inter-vehicle distance is derived. To single out the delay effect, we impose the following assumption.

(1) Direct Transmission of Braking Action

Suppose that the leading vehicle transmits its braking action directly to the vehicle behind it. This is the fastest way to inform the following vehicle to take action. If no time delay is involved, then the following vehicle will brake immediately and the inter-vehicle distance will be kept contact until both vehicles come to the complete stop. However, communication delays will postpone the following vehicle's action. The main question is: How much delay can be tolerated?

Assumption 2: (1) The leading vehicle and following vehicle travel at the cruising condition with distance d_{ref} and speed $v(0)$. (2) The information on the braking action $F_0 = -F_{max}$ of the leading vehicle at $t = 0$ is immediately transmitted to vehicle 1 with a communication delay τ. (3) No other information is available to vehicle 1.

Theorem 3: Under Assumption 2, the final distance d_1^{final} is

$$d_1^{final} = d_{ref} - v(0)\tau + \frac{F_{max}}{2m}\tau^2.$$

Proof: Since the braking force for the leading vehicle is $-F_{max}$, its speed profile is

$$v_0(t) = v(0) - \frac{F_{max}}{m} t.$$

At time τ, its speed is

$$v_0(\tau) = v(0) - \frac{F_{max}}{m} \tau.$$

Vehicle 1 receives the braking information at τ and immediately applies the maximum braking force $-F_{max}$ with the initial speed $v(0)$. As a result, $\eta = v^2(0) - v_0^2(\tau)$.

By Theorem 1, the final distance is

$$\begin{aligned} d_1^{final} &= d_{ref} - \frac{\eta m}{2F_{max}} \\ &= d_{ref} - \frac{(v^2(0) - v_0^2(\tau))m}{2F_{max}} \\ &= d_{ref} - \frac{(v^2(0) - (v(0) - \frac{F_{max}}{m}\tau)^2)m}{2F_{max}} \\ &= d_{ref} - \frac{(2v(0)\frac{F_{max}}{m}\tau - \frac{F_{max}^2}{m^2}\tau^2)m}{2F_{max}} \\ &= d_{ref} - v(0)\tau + \frac{F_{max}}{2m}\tau^2. \end{aligned}$$

□

Corollary 1: For a given required minimum distance d_{min}, the maximum tolerable communication delay is

$$\tau_{max} = \frac{v(0) - \sqrt{v^2(0) - 2\frac{F_{max}}{m}(d_{ref} - d_{min})}}{2}.$$

Proof: By Theorem 3, to satisfy $d_1^{final} \geq d_{min}$, the maximum tolerable τ is solved from $d_{min} = d_{ref} - v(0)\tau + \frac{F_{max}}{2m}\tau^2$ or

$$\frac{F_{max}}{2m}\tau^2 - v(0)\tau + (d_{ref} - d_{min}) = 0$$

whose smaller solution is

$$\tau_{max} = \frac{v(0) - \sqrt{v^2(0) - 2\frac{F_{max}}{m}(d_{ref} - d_{min})}}{2}.$$

□

In particular, for collision avoidance, $d_{min} = 0$ and we have

$$\tau_{max} = \frac{v(0) - \sqrt{v^2(0) - 2\frac{F_{max}}{m}d_{ref}}}{2}.$$

For the evaluation scenario in Section III-B, $v(0) = 25$, $F_{max} = 10000$, $m = 1500$, $d_{ref} = 40$, and $d_{min} = 15$. The corresponding maximum tolerable delay is $\tau_{max} = 3.9609$ second. For collision avoidance, $d_{min} = 0$ and $\tau_{max} = 7.7129$ second.

However, if the vehicle weight is increased to $m = 1800$ (Kg) and the platoon cruising distance is reduced to $d_{ref} = 30$, the tolerable delay is reduced to $\tau_{max} = 1.7956$ second.

Typical vehicle braking control must balance safety and driveability. Consequently, inter-vehicle distances may reduce more significantly than the scenario of this subsection. As a result, the maximum tolerable delay may be significantly less. These will be evaluated in the subsequent case studies.

(2) Broadcasting Schemes and Consequence

The leading vehicle's braking action can be broadcasted to the platoon. The average communication latency depends on the distance between the sending (leading vehicle) node and the receiving node. Using the basic square relationship, if the first following vehicle experiences a delay $\tau_1 = \tau$, then the second vehicle will have a delay around $\tau_2 = 4\tau$, the third vehicle with $\tau_3 = 9\tau$, and so on.

For example, if $d_{ref} = 40$ (m) and $\tau_1 = 100$ (ms), then $\tau_2 = 400$ (ms) (at 80 (m)), ..., $\tau_7 = 4.9$ (s) (at 280 (m)), which implies that d_7 will fall below 15 m, violating the minimum distance requirement.

This analysis indicates that communication schemes need to be carefully designed when a platoon has many vehicles.

V. Communication Systems

A. Communication Standards and Latency

To study more realistically how communication systems and control interact, we use a generic communication scheme shown in Fig. 9. In this scheme, a data packet is generated and enters the queue for transmission. The queuing time depends on network traffic and data priorities. The packet contains both data bits and error checking bits. We assume that the error checking mechanism is sufficient to detect any faulty packet. If the packet transmission is successful, the receiver returns an acknowledgment message to the sender, which completes the transmission. If the packet is received with error, it will be discarded and a request is sent back to the sender to re-transmit the same packet. The permitted total time for transmission of a packet is pre-determined by the control updating times. If a packet was not successfully transmitted when the control updating time is up, the packet will be considered as lost.

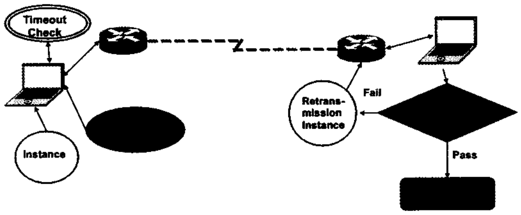


Fig. 9. Data transmission schemes.

Inter-vehicle communications (IVC) can be realized by using infrared, radio, or microwaves waves. For instance, in IEEE 802.11p, a bandwidth 75 MHz is allotted in the 5.9 GHz band for dedicated short range communication (DSRC) [19], [20]. Alternatively, ultra-wideband (UWB) technologies have been used for IVC. IEEE 802.11x, where $x \in \{a, b, g, p \ldots\}$ have been studied for inter-vehicle use. At present, many applications use DSRC with IEEE 802.11p (a modified version of IEEE 802.11 (WIFI) standard) at the PHY and MAC layers. IEEE 802.11g and IEEE 802.11p are used for experimental studies in this paper.

In the middle of protocol stack, DSRC employs IEEE 1609.4 for channel switching, 1609.3 for network service, and 1609.2 for security service. In the network service, users have a choice between the wireless access for vehicle environments short message protocol (WSMP) or the internet protocol version 6 (IPv6) and user datagram protocol (UDP)/transmission control protocol (TCP). Single-hop messages typically use the bandwidth-efficient WSMP, while multi-hop packets use the IPv6+UPD/TCP for its routing capability.

Inter-vehicle communications use wireless networks that are subject to severe uncertainties. For example, the signal-to-interference-plus-noise ratio (SINR) [21] attenuates with distance (it decreases inverse proportionally to the cubic of the distance between the two vehicles). It is also affected by obstructions such as buildings, bridges, other vehicles, etc. Other factors include queue delays, network data traffic conditions, routes, signal fading, signal interference from other vehicles, Doppler shifts, and traffic and weather conditions. These uncertainties depend significantly on channel coding schemes and communication networks. These factors collectively determine packet delivery delays, packet loss rates, etc. This paper will focus on delay effects. To be concrete in treating communication systems, we will employ IEEE 802.11 standards as our benchmark systems and the related latency data [19].

Bandwidth-delay product is often used to characterize the ability of a network pathway in carrying data flows [22], [23]. When the TCP protocol is used in data communications, packet-carrying capacity of a path between two vehicles will be limited by this product's upper bound. For more detailed discussions on capacity/delay tradeoffs, the reader is referred to [19] and the references therein. Note that latency is further caused by delays in each hub's queues, routes (multi-hub), packet delivery round-trip time, channel reliability, retransmission, scheduling policies in interference avoidance strategies. Although typical transmission delays can be as low as several millisecond, vehicular traffic scenarios introduce combined latency of several hundreds of milliseconds even several seconds. In this paper, we will show that delays of such scales will have significant impact on vehicle safety.

B. A Single-Hop Experimental Study

We assume the three-vehicle scenario in Fig. 2. Communication channels between v_0 and v_2 use the WSMP protocol. This protocol can carry messages on both the Control Channel (CCH) and the Service Channel (SCH). The WSMP allows direct control of the lower-layer parameters such as transmission power, data rates, channel numbers, and receiver MAC addresses. The WSMP over the CCH can skip the steps of forming a WAVE Basic Service Set (BSS) that delivers IP and WAVE short message (WSM) data on the SCH. Those methods can potentially reduce communication latency.

The round trip time (RTT) under this protocol includes measurement time for the variables (vehicle distance, speed, etc.), source data creation time (creating packets, adding verification codes, scheduling, etc), communicating the packet to v_2, receiver verification, travel time for sending back acknowledgment from v_2. Fig. 10 sketches some of the time delays from these steps.

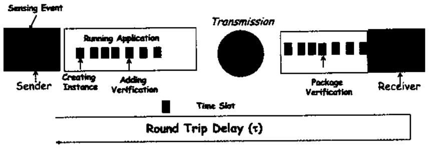


Fig. 10. A Round Trip Delay.

In an ideal case that v_0 can capture the CCH during each CCH time slot, v_0 can send its beacon and update its status to v_2 at the rate of 10 Hz. If a package is successfully transmitted and verified during the first round, the Package Delivery Rate(PDR) is 1, the RTT $\tau^0 \leq 100$ ms since IEEE 1609.4 specifies the reoccurrence of the CCH at the rate of every 100 ms.

The physical limitations on wireless channels (bandwidth and power constraints, multi-path fading, noise and interference) present a fundamental technical challenge to reliable high-speed communication. One or several retransmissions are often necessary to meet a PDR requirement. In this case, delay is $\tau = n\tau^0$ where n is the number of average rounds for a successful transmission. In the following examples, we show how modulation rates and channel interferences affect the number of retransmission and delay τ. Due to the network system heterogeneity and highway environments, we are using the truth-ground data, rather than ns-3 simulations.

Example 1: [25] reports experimental data of IEEE 802.11p DSRC from a team of vehicles driving on certain Michigan highways. Package Delivery Rates (PDR) are measured under different driving conditions, traffics, and surroundings. A typical curve from [25] is re-generated in Fig. 11. When the modulation rate is 6 Mbps, the Package Delivery Rate (PDR) is about 75% at a distance of 85 m. The first round-trip takes about 100 ms. Each subsequent round-trip must catch the next CCH and it takes on average more than three retransmissions to achieve a PDR over 98.5%. Consequently, the average delay is $\tau \approx 0.3$ second. When the modulation rate is increased to 18 Mbps, the PDR is reduced to 36% at 85 m. In order to meet the same PDR 98.5%, the delay is more than 1 second.

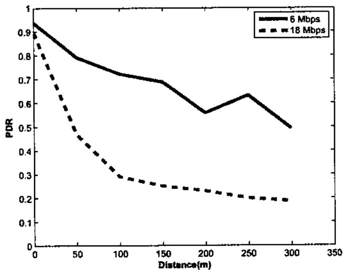


Fig. 11. PDR vs. separation distance under different data rates in the Rural Road (RR) environment (with 95% Confidence Interval). Here, the data rates are 6 Mbps and 18 Mbps. The transmission power is 20 dBm.

Example 2: In this experimental study, we use the IEEE 802.11g standard to analyze the affects of multi-path interference. The communicating nodes reside on laptop computers and are moved from a short distance of 20 m to 95 m. In the first experimental setting, the transmission pathway does not have obvious obstacles, except low grass on the open field. Communication latency is recorded by the synchronized clocks on these computers. Fig. 12 provides the experiment data on recorded latency for different inter-node distances. A simplified curve can be obtained by data fitting, which is also shown in the same figure. It is noted that latency between 100 ms to 600 ms is typical in this case study.

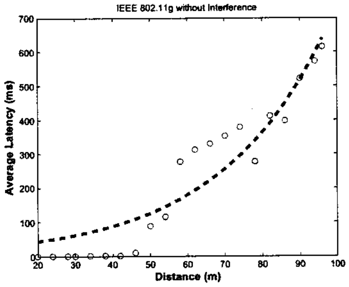


Fig. 12. Dependence of latency on distance without obstacles on the transmission pathway.

Example 3: Extending on the experiment in Example 2, we now evaluate impact of obstacles on transmission pathways. Under the same experimental protocols as in Example 2, we select a field with many trees, but not overly dense. Consequently, depending on distances, the transmission pathways are obstructed by several trees. Fig. 13 demonstrates the experimental data on communication latency under different transmission distances. It is seen clearly that with obstacles, communication latency increases significantly to a range of 3.4 second.

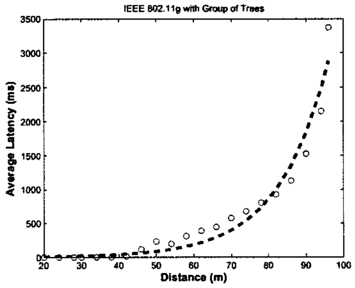


Fig. 13. Dependence of latency on distance with trees on the transmission pathway.

C. Multi-Hop Communication Data

Inter-vehicle communications may involve multi-hops which create further delays. Typically, the IPv6+UDP/TCP protocols can be used in such systems. Unlike the WSMP protocols which use 11 bytes overhead, the IPv6 protocol requires a minimum overhead of 52 bytes. Although this is more complicated in coding and less efficient in using the data resource, this protocol provides more flexible routing schemes. There are many experimental studies of IEEE 802.11p under multi-hop and highway environment. Since we are only concerned with latency data, we quote here the studies in [19] that contain extensive experimental results. A typical curve from [19] is re-generated in Fig. 14. It is noted that although IEEE 802.11p uses higher power and faster speed, a latency of hundreds of milliseconds is typical in highway conditions.

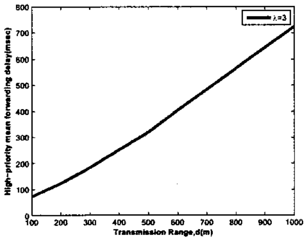


Fig. 14. Average delay of high-priority message dissemination for 5 hops of communication as functions of the transmission range.

VI. Platoon Information Structure

A. Safety under Front Sensor Information

We start with the basic information structure of using front distance sensors only. For the three-car platoon in Fig. 2 and the control law $F = g_1(d)$ in Fig. 4, the closed-loop system becomes

$$\left\{ \begin{array}{rcl} \dot{v}_0 & = & \frac{1}{m_0}(F_0 - (a_0 v_0 + b_0 v_0^2)) \\ \dot{v}_1 & = & \frac{1}{m_1}(g_1(d_1) - (a_1 v_1 + b_1 v_1^2)) \\ \dot{v}_2 & = & \frac{1}{m_2}(g_1(d_2) - (a_2 v_2 + b_2 v_2^2)) \\ \dot{d}_1 & = & v_0 - v_1 \\ \dot{d}_2 & = & v_1 - v_2 \end{array} \right. \tag{8}$$

Example 4: We consider the scenario defined in Section III-B. Suppose that the platoon uses only front sensors to measure inter-vehicle distances, namely the information structure (a) in Fig. 2 is in effect. The feedback control function $F = g_1(d)$ is depicted in Fig. 6. We will use the following fast braking condition for comparison.

Under the **Fast Braking** scenario from Section III-B, suppose that the leading vehicle uses a braking force 5000 N, which brings it to a stop from 25 m/s in 7.5 second. The distance trajectories of d_1 and d_2 are shown in Fig. 15. In this case, the minimum distances are 20.6 m for d_1 that is acceptable, but 0 m for d_2. This means that vehicle 2 will collide with vehicle 1 during the transient time.

To explain this scenario, we note in the top plot of Fig. 15 that since vehicle 2 relies on d_2 to exercise its braking control function, there is a dynamic delay in initiating its braking. d_2 is reduced to about 20 m when vehicle 2 starts to act. For a large platoon, this dynamic delay from vehicle to vehicle is a serious safety concern.

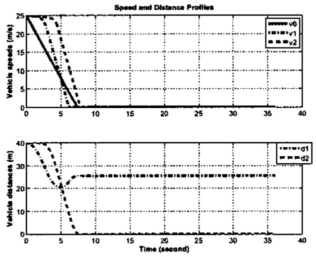


Fig. 15. Distance trajectories under fast braking.

B. Adding Distance Information by Communications

We next expand on the information structures beyond front sensors by adding distance information by communications.

Example 5: Continuing from Example 4, we consider the same three-car platoon under the same initial conditions: The nominal inter-vehicle distances are 40 m; the cruising vehicle speeds are 25 m/s; the maximum braking force is 10000 N.

Under the **Fast Braking** scenario as in Example 4, suppose now that vehicle 1 sends d_1 information to vehicle 2 by communication. As a result, vehicle 2 can use both d_1 and d_2 in its control function; see Fig. 16.

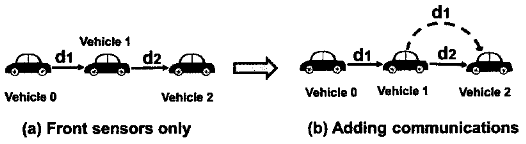


Fig. 16. Enhanced information structure by sending d_1 to vehicle 2 by communication links in Example 5

Suppose that vehicle 2 modifies its braking control function from the previous $F_2 = g_1(d_2)$ to the weighted sum $F_2 = 0.5g_1(d_2) + 0.5g_1(d_1)$ that uses both distances. The resulting speed and distance trajectories are displayed in Fig. 17. Now, the minimum distances are 20.6 m for d_1 and 15.9 m for d_2, both are within the safety region.

To compare Fig. 15 and Fig. 17, we note that with information feeding of d_1 into vehicle 2, vehicle 2 can slow down when d_1 is reducing before d_2 changes. Consequently, it is able to act earlier, resulting in a reduced distance swing for d_2 during the transient.

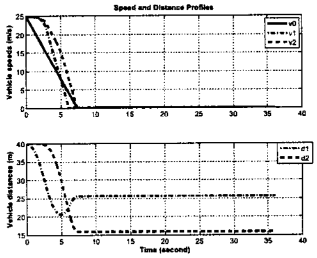


Fig. 17. Distance trajectories when the distance information d_1 is made available to vehicle 2. It shows improvement over Fig. 15.

VII. Platoon Information Contents

A. Adding Speed Information by Communications

We now add the speed information of the leading vehicle to both vehicles 1 and 2 by communication.

Example 6: For the same three-car platoon under the same initial conditions as Example 5, we add the leading vehicle's speed v_0 into the information structure. This information is transmitted (or broadcasted) to both vehicles 1 and 2. Under the **Fast Braking** scenario as in Example5, suppose that vehicles 1 and 2 receive the additional speed information v_0, resulting in a new information structure shown in Fig. 18.

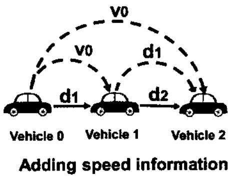


Fig. 18. Enhanced information structure by sending d_1 to vehicle 2 and v_0 to both vehicles 1 and 2.

From the control functions of Example 5, additional control actions $g_2(v_0, v_1)$ and $g_2(v_0, v_2)$ are inserted. The resulting speed and distance trajectories are displayed in Fig. 19. Now, the minimum distances are 28.3 m for d_1 and 27.1 m for d_2, a much improved safety performance.

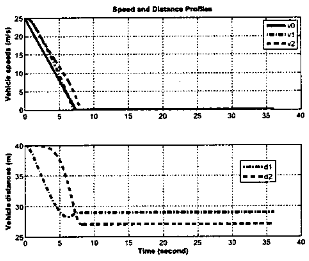


Fig. 19. Distance trajectories when both distance and speed information is made available.

B. Adding Braking Event Information by Communications

Intuitively, if the leading vehicle's braking action can also be communicated, the following vehicles can act much earlier than their measurement data on vehicle movements. To evaluate benefits of sending the driver's action, we add the braking event information of the leading vehicle to vehicle 2 by communications.

Example 7: For the same three-car platoon under the same initial conditions as Example 6, we now further add the leading vehicle's braking event information F_0 into the information structure. To understand the impact, we purposely assume that vehicle 1 does not receive this information. In other words, this information will be transmitted only to vehicle 2 by communications. Under the **Fast Braking** scenario as in Example 6, suppose that vehicle 2 receives the additional braking event information F_0, resulting in a new information structure shown in Fig. 20.

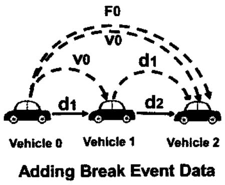


Fig. 20. Enhanced information structure by sending the braking event F_0 to vehicle 2.

From the control functions of Example 6, an alternative control action F_0 is inserted when $d_2 < d_{ref} = 40$ m. The resulting speed and distance trajectories are displayed in Fig. 21. Now, the minimum distances are 28.3 m for d_1 and 30.6 m for d_2, a much improved safety over the case in Example 6. It is interesting to note that by knowing the leading vehicle's action, vehicle 2 can react faster than even vehicle 1 which does not receive the braking action data.

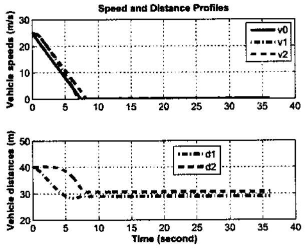


Fig. 21. Distance trajectories with added braking event information.

VIII. Impact of Radar and Communication Uncertainties

A. Impact of Radar Resolution and Missed Detection

Radar sensors provide a stream of measurement data, typically using 24, 35, 76.5, and 79 GHz radars. In general, radar sensor measurements are influenced by many factors that limit their accuracy and reliability. These include signal attenuation by the medium, beam dispersion, noises, interference, multi-object echo (clutter), jamming, etc.

We first consider the impact of radar's resolution on a platoon system. Within the same setup as Example 5, vehicle 2 receives the distance information of d_1 and d_2 in which d_2 is measured by a radar. Taking into consideration radar resolution, the measured distance is $\tilde{d}_2 = d_2 + \gamma\delta$, where γ is a resolution level and δ is a standard Gaussian noise $\mathcal{N}(0, 1)$.

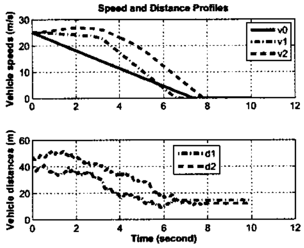


Fig. 22. Distance trajectories under a radar of low resolution (1 m).

Fig. 22 shows a simulation result under a radar of resolution 1 m. The distribution of the minimum distances after repeated runs to account for randomness is shown in Fig. 23. Although the expectation is 8.01 m, the minimum distance has a high probability of having values close to zero. Consequently, this low resolution radar is not suitable for this application.

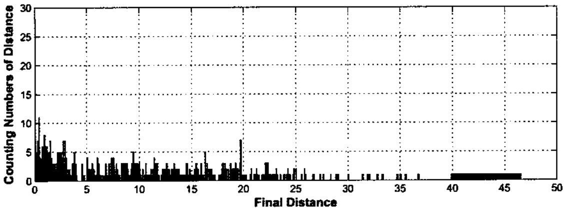


Fig. 23. The distribution of minimum distances d_2 under a radar of low resolution (1 m).

Next, we upgrade the radar to a higher resolution 0.1 m. A corresponding simulation is shown at Fig. 24. The minimum distances for both d_1 and d_2 are much improved. The distribution of minimum distances of d_2 is shown in Fig. 25. The random minimum distances have expectation 15.92 m and variance $\sigma^2 = 0.31$. This is an acceptable resolution for this application.

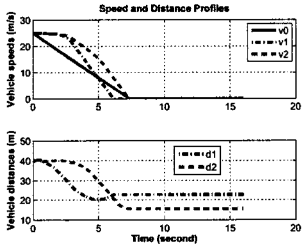


Fig. 24. Distance trajectories with high Resolution Radar.

It is noted further that uncertainties of radar signals include also random false alarms or missed detection. In this scenario, the sensor does not provide information at the sampling time, and the control/brake action must rely on its previous measurements and other available information from different resources. This situation is similar to Example 12 when communication information is unavailable, which will be detailed in the next subsection.

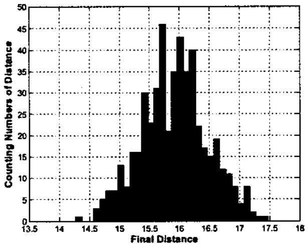


Fig. 25. Distance Distribution of d_2 with high Resolution Radar .

B. Impact of Communication Delay

Communications introduce a variety of uncertainties. Most common types are communication latency and packet loss. These can be caused by many factors as listed in Section I. This paper focuses on communication latency. Depending on environment and communication protocols, communication latency can be near a constant, distance dependent, or random. We cover these cases in the following subsections.

1) Fixed Delays: We first consider fixed delays.

Example 8: Under the same system and operating condition as Example 5, we assume that the communication channel for the distance information has a delay of τ second. The impact of the communication delay is shown in Fig. 26. Without the delay, the minimum distance for d_2 is 15.9 m. When a delay of $\tau = 0.6$ (second) is introduced, the minimum distance for d_2 is reduced to 11 m.

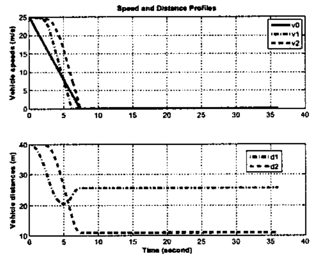


Fig. 26. Distance trajectories when communication delays are considered.

Table I lists the relationship between the delay time and the minimum distance for d_2.

TABLE I
IMPACT OF COMMUNICATION DELAYS

delay time τ (s)	0	0.3	0.6	0.9	1.2
minimum d_2 (m)	15.9	13.6	11	8.2	5.1

Next, we use experimental delay data in our simulation studies.

Example 9: Under the same system and operating condition as Example 5, we assume that communication systems use the single-hop scenario in Section V-B. Under a scenario of latency $\tau = 0.1$ second (CCH delay only), the minimum distance for d_2 is 15.1 m. It remains as an acceptable safe distance. Many factors affect such delays. One essential consideration is channel capacity. Shannon's channel capacity claims that if the channel is too noisy which reduces channel capacity, information cannot be effectively transmitted. This is translated into very large channel latency under a required PDR. In this sense, impact analysis of channel latency is in fact a study on communication resources. Here we use platoon safety as a performance criterion in this study.

2) Distance-Dependent Delays: In vehicle platoon environment, communication latency depends directly on inter-vehicle distances. These are reflected clearly in Figures 11, 12, and 13. It is observed that during platoon formation and braking, inter-vehicle distances change substantially. This subsection considers delays as a function of distance.

Example 10: Under the same system and operating condition as Example 9, we now use more realistic experimental data in Fig. 12 for latency which is a function of distance. Based on the relationship of distance and latency, the simulation in Fig. 27 shows that the minimum distance for d_2 is now 12.7 m. Furthermore, if signal interference, obstructions, and fading are considered, the latency is increased to these in Fig. 13. The simulation results in a minimum distance for d_2 as 5.6 m. This is shown in Fig. 28, which causes safety concerns.

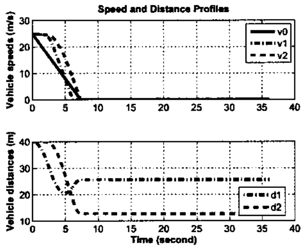


Fig. 27. Distance trajectories when communication delays are dependent on vehicle distances, whose function form is given in Fig. 12 for the "no obstacle" scenario.

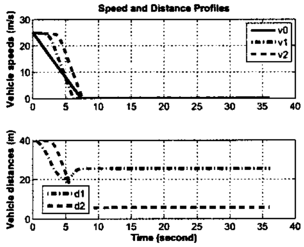


Fig. 28. Distance trajectories when communication pathways are obstructed as shown at Fig. 13.

Example 11: Continuing the study of Example 9, we consider the multi-hop scenario in Subsection V-C. In that scenario, transmission from v_0 to v_2 is over 5 hops. Suppose that each hop has the same priority, and that each loses CCH once followed by one successful re-transmission. Based on the distances between the vehicles in the example, the total communication delay $\tau > 1.5$ second. The simulation shows that the minimum distance for d_2 approaches to 0, leading to a collision.

3) Random Delays: Typically, communication delays are random variables with certain distributions. Depending on latency control mechanisms of transmission protocols, the latency can have different distributions. We use the common Gaussian distribution for our study in this subsection.

Example 12: Assume that communication latency is a random variable, due to the random features of wireless transmissions. In this example, we model τ as a random variable that is Gaussian distributed with $\mathbb{E}(\tau) = 1.2$ (second) and variance $\sigma^2 = 0.09$. Continuing the study of Example 11, the simulation in Fig. 29 shows that the minimum distance d_2 approaches to 5.09 m.

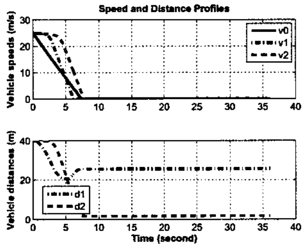


Fig. 29. Distance trajectories under communication latency which is Gaussian distributed.

Simulation results of minimum distance distribution are shown at Fig. 30. The variance of d_2 is $\sigma^2 = 0.142$.

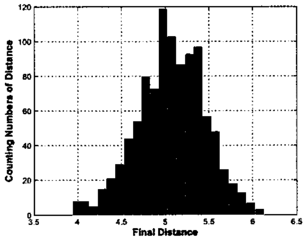


Fig. 30. Distance distribution of d_2 under random communication latency .

C. Impact of Doppler Frequency Shift and Signal Spreading

Mobility-induced Doppler spread is one of the main factors that degrade the performance of Orthogonal Frequency Division Multiplexing (OFDM) schemes. It introduces Inter-Symbol Interference (ISI) and Inter-Carrier Interference (ICI) by destroying the orthogonality between adjacent sub-carriers.

In most cases, DSRC is adequate in restoring both zero ISI and zero ICI in highly mobile, severe-fading vehicular environments, as discussed with great detail in [24]. In the physical layer of IEEE 802.11p, the bandwidth of each DSRC channel is 10 MHz, which entails less ISI and ICI than IEEE

802.11a which uses 20 MHz channel bandwidth. This brings better wireless channel propagation with respect to multi-path delay spreads and Doppler effects caused by high mobility and roadway environments. Also, DSRC expands Guard Band (GB) to 156 KHz and has $1.6\mu s$ guard interval for OFDM schemes. The Guard Band between sub-carriers can ensure that mobility-induced Doppler spreads do not cause two adjacent sub-carriers to overlap.

On the other hand, with high operation frequency at 5.9 GHz, IEEE 802.11p is subject to higher Doppler frequency shifts. When vehicle speeds are extremely high (such as 250 km/h on German highways), the issue of Doppler frequency shifts become more pronounced. At present, fast network topology switching and complicated road environments are still challenges with respect of ISI and ICI, and remain to be resolved by new technologies.

Fig. 31, re-produced from [25], compares the impacts of Open Field (OF) and Rural Freeway (RRF) on the PDR. The PDR remains nearly unchanged in the OF environment when relative vehicle velocities vary from 0 (m/s) to 25 (m/s). In contrast, the PDR drops dramatically in the RRF environment. For example, when the relative velocity is 12.5 (m/s), the PDR of the communication link in the RRF environment is reduced to $1/3$ of that with the OF environment. This implies that in the RRF environment, much more communication resources are needed to ensure the same level of safety. As a result, it is advisable that these DSRC characteristics be incorporated into the platoon design by VANET designers.

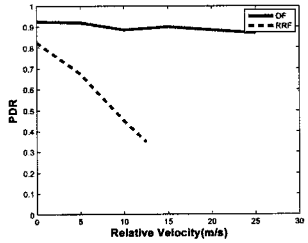


Fig. 31. The impact of relative velocities on the PDR(with the 95% confidence interval). A bin of 20 packets is used to calculate PDR values as well as relative velocities.

D. System Integration with VANET Framework

The generic platoon model of this paper is an important component of a VANET framework as shown in Fig. 32. In our exploration, the actual communication routes are not specified. Within a VANET, the links among vehicles can be realized by V2V communications or V2I pathways involving access points, wireless towers and other infrastructures. Our model provides a fundamental framework to study impact of communications on vehicle safety and can be specified to different communication configurations. The findings of this paper can be used as guidelines in selecting VANET parameters. For example, transmission power, modulation rate, and coding scheme can be selected so that they meet the requirements of an acceptable minimum inter-vehicle distance. Also, a platoon can potentially enhance VANET data access performance. By using vehicles as transmission hubs, data can be replicated and relayed to more vehicles in the group. This structure improves VANET resources in a distributed manner and, if used properly, can improve overall performance.

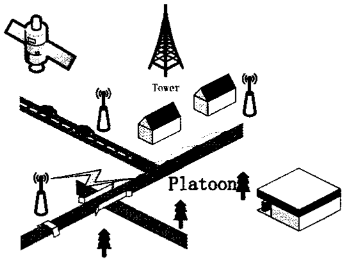


Fig. 32. System integration of a platoon with a VANET framework.

While this paper is focused on one platoon formation, a platoon experiences many dynamic variations in real implementations. These include lane change, vehicle departure and addition, platoon reformation, etc. At the network level, such changes amount to network topology variations. At the communication/physical level, some uncertainties will be introduced such as echo among vehicles and road infrastructures. A VANET can easily accommodate such topology changes by using vehicle IDs and their links. Furthermore, by seamless integration into a VANET, a platoon can have access to VANET resourses, including GPS, Internet, distributed live database, VANET-enabled applications, etc. Consequently, a platoon can potentially utilize additional information in its safety considerations via inter-vehicle communications and emission reduction via traffic information. These topics are, however, beyond the scope of this paper. We refer the reader to [26], [27] for some related studies.

IX. Discussions and Concluding Remarks

This paper investigates the interaction between control and communications, in the framework of highway platoon safety. Information structure, information content, and information reliability have been taken into consideration in this study. It is well perceived that communication systems introduce uncertainties that are of many types and values. To be concrete, we have selected communication latency as a key uncertainty in this study.

The main results of this paper demonstrate that communications provide critical information that can enhance vehicle safety effectively beyond distance sensors. In fact, from our simulation studies, platoon control may mandate communications for additional information. Although traditionally, distance and vehicle speed are immediate candidates for transmission, our results show that drivers' braking events contain very effective information for platoon management. Our simulations suggest that platoon communications place event data under more prominent considerations.

Our study shows that communication latency is a critical factor in information exchange. Large latency can diminish values of data communication in platoon control. It is a common framework in multi-vehicle communication scenarios that vehicles within an interference radius do not transmit simultaneously. A direct consequence is that latency becomes larger. For instance, under the IEEE 802.11p standard, transmission radius can reach 1 km. If 50 vehicles are in this region and each transmission (or broadcasting) takes 30 ms, a delay of 1.5 second will occur between consecutive transmissions of a given vehicle. Our study shows that such a delay has an alarmingly high impact on vehicle safety. This issue deserves further studies.

What is reported in this paper is a first step in this direction. There are many un-resolved issues. We are currently investigating the impact of communications on platoon safety under packet erasure channels. Furthermore, we have only considered basic driving conditions: Straight lanes, dry surface conditions, good weather conditions, and no lane changes or platoon re-formation after vehicle departure or addition. System integration with VANET framework is a worthy topic to pursue. All these issues are worth further studies.

REFERENCES

[1] J.K. Hedrick, D. McMahon, D. Swaroop, Vehicle modeling and control for automated highway systems, PATH Research Report, UCB-ITS-PRR-93-24, 1993.

[2] P. Ioannou and C. Chien, Autonomous intelligent cruise control, *IEEE Transactions on Vehicular Technology*, Vol. 42, No. 4, pp. 657672, 1993.

[3] R. Rajamani, H.S. Tan, B. Law and W.B. Zhang, Demonstration of integrated lateral and longitudinal control for the operation of automated vehicles in platoons, *IEEE Transactions on Control Systems Technology*, Vol. 8, No. 4, pp. 695-708, 2000.

[4] F. Knorr, D. Baselt, M. Schreckenberg, and M. Mauve, Reducing traffic jams via VANETs, *IEEE Transactions on Vehicular Technology*, Vol. 61 , Iss. 8, pp. 3490-3498, 2012.

[5] K.S. Chang, W. Li, P. Devlin, A. Shaikhbahai, P. Varaiya, J.K. Hedrick, D. McMahon, V. Narendran, D. Swaroop, J. Olds, Experimentation with a vehicle platoon control system, Vehicle Navigation and Information Systems Conference, pp. 1117-1124, 1991.

[6] D.H. Narendran, V.K. Swaroop, D. Hedrick, J.K. Chang, K.S. Devlin, P.E., Longitudinal Vehicle Controllers for IVHS: Theory and Experiment McMahon,American Control Conference, 1992 Publication Year: 1992 , Page(s): 1753 - 1757

[7] Y.F. Zhao, H. Ogai, Development of a platooning control algorithm based on RoboCar, ICE Annual Conference (SICE), pp. 352-355, 2011.

[8] J. Bom, B. Thuilot, F. Marmoiton, P. Martinet, A global control strategy for urban vehicles platooning relying on nonlinear decoupling laws, Intelligent Robots and Systems, pp. 2875-2880, 2005.

[9] G. Guo and W. Yue, Autonomous platoon control allowing range-limited sensors, *IEEE Transactions on Vehicular Technology*, Vol. 61, Iss. 7, pp. 2901-2912, 2012.

[10] C.Y. Liang and H. Peng, String stability analysis of adaptive cruise controlled vehicles, *JSME International Journal Series C*, Vol. 43, Iss. 3, pp. 671-677, 2000.

[11] L.Y. Wang, A. Syed, G. Yin, A. Pandya, H.W. Zhang, Coordinated vehicle platoon control: weighted and constrained consensus and communication network topologies, *Proceedings of CDC 2012*, Hawaii, pp. 4057-4062, Dec. 2012.

[12] L.Y. Wang, A. Syed, G. Yin, A. Pandya, H.W. Zhang, Control of vehicle platoons for highway safety and efficient utility: Consensus with communications and vehicle dynamics, *Journal of Systems Science and Complexity*, accepted and to appear in 2013.

[13] J. S. Freudenberg and R. H. Middleton, Feedback control performance over a noisy communication channel. *Proceedings of the 2008 Information Theory Workshop*, Porto, Portugal, pp. 232-236, May 2008.

[14] R. Luck and A. Ray, Experimental verification of a delay compensation algorithm for integrated communication and control system. *International Journal of Control*, vol. 59, pp. 1357-1372, 1994.

[15] A. J. Rojas, J. H. Braslavsky, and R. H, Middleton, Fundamental limitations in control over a communication channel *Automatica*, Vol. 44, pp. 3147-3151, 2008.

[16] D.H. McMahon, J.K. Hedrick, S.E. Shladover, Vehicle modelling and control for automated highway systems, *Proceedings of American Control Conference*, San Diego, CA, USA, pp. 297 - 303, May 23-25, 1990.

[17] R.T. O'Brien, Vehicle lateral control for automated highway systems, *IEEE Transactions on Control Systems Technology*, Vol. 4, Iss. 3, pp. 266 - 273, 1996.

[18] S. Sheikholeslam, C.A. Desoer, Longitudinal control of a platoon of vehicles with no communication of lead vehicle information: a system level study, *IEEE Transactions on Vehicular Technology*, Vol. 42, Iss. 4, pp. 546 - 554, 1993.

[19] M.J. Neely and E. Modiano, Capacity and delay tradeoffs for Ad-Hoc mobile networks, *IEEE Tran. on Information Theory*, Vol. 51, No. 6, pp. 1917-1936, 2005.

[20] K.A. Hafeez, L. Zhao, B. Ma, J.W. Mark, Performance analysis and enhancement of the DSRC for VANET's safety applications, *IEEE Transactions on Vehicular Technology*, Vol. 62 , Iss. 7, pp. 3069-3083, 2013.

[21] A. Kumar, D. Manjunath, J. Kuri, Wireless Networking, Elsevier, 2008.

[22] L.L. Peterson, B.S. Davie, Computer Networks (2nd Ed.), Morgan Kaufmann, San Francisco, CA, USA, 2000.

[23] W. Stevens, TCP/IP Illustrated (Vol. 1, The Protocols), Addison-Wesley, Reading, MA, USA, 1994.

[24] L. Cheng, B.E. Henry, D.D. Stancil, F. Bai, and P. Mudalige, Mobile vehicle-to-vehicle narrow-band channel measurement and characterization of 5.9 GHz dedicated short range communication (DSRC) frequency band, *IEEE Journal on Selected Areas in Communication*, Vol. 25, Issue 8, pp. 1501-1516, October 2007.

[25] F. Bai, D.D. Stancil, H. Krishnan, Toward understanding characteristics of dedicated short range communications (DSRC) from a perspective of vehicular network engineers, *Proceedings of MobiCom10*, Chicargo, IL, USA, pp. 329 - 340, September 20-24, 2010.

[26] Le Yi Wang, Ali Syed, George Yin, Abhilash Pandya, Hongwei Zhang, Control of vehicle platoons for highway safety and efficient utility: Consensus with communications and vehicle dynamics, *Journal of Systems Science and Complexity*, accepted and to appear.

[27] Le Yi Wang, Ali Syed, George Yin, Abhilash Pandya, Hongwei Zhang, Coordinated vehicle platoon control: weighted and constrained consensus and communication network topologies, *Proceedings of CDC 2012*, Hawaii, pp. 4057-4062, Dec. 2012.

Lijian Xu (S'12) received B.S. from University of Science and Technology Beijing China in 1998, M.S. from University of Central Florida in 2001, then he worked for AT&T, FL and Telus Communication, Canada as an engineer and engineering manager from 2001 to 2007. Now he is a PhD candidate in Electrical and Computer Engineering Department at Wayne State University. His interests are in the areas of network control system, digital wireless communication, consensus control, vehicle platoon and safety. He received The Best Paper Award from 2012 IEEE International Conference on Electro/Information Technology. He is a student member of IEEE.

Le Yi Wang (S'85-M'89-SM'01-F'12) received the Ph.D. degree in electrical engineering from McGill University, Montreal, Canada, in 1990. Since 1990, he has been with Wayne State University, Detroit, Michigan, where he is currently a Professor in the Department of Electrical and Computer Engineering. His research interests are in the areas of complexity and information, system identification, robust control, H^∞ optimization, time-varying systems, adaptive systems, hybrid and nonlinear systems, information processing and learning, as well as medical, automotive, communications, power systems, and computer applications of control methodologies. He was a keynote speaker in several international conferences. He was an Associate Editor of the IEEE Transactions on Automatic Control and several other journals, and currently is an Associate Editor of the Journal of System Sciences and Complexity and Journal of Control Theory and Applications. He is a Fellow of IEEE.

George G. Yin (S'87-M'87-SM'96-F'02) received the B.S. degree in mathematics from the University of Delaware in 1983, M.S. in Electrical Engineering, and Ph.D. in Applied Mathematics from Brown University in 1987. He joined Wayne State University in 1987, and became a professor in 1996. His research interests include stochastic systems, applied stochastic processes and applications. He severed on many technical committees; was the Co-chair of a couple of AMS-IMS-SIAM Summer Conferences, and the Co-chair of 2011 SIAM Control Conference. He is or was an associate editor of many journals including SIAM Journal on Control and Optimization, Automatica, and IEEE Transactions on Automatic Control. He is a Fellow of IEEE.

Hongwei Zhang (S'01-M'07-SM'13) received his B.S. and M.S. degrees in Computer Engineering from Chongqing University, China and his Ph.D. degree in Computer Science and Engineering from The Ohio State University, USA. He is currently an associate professor of computer science at Wayne State University. His primary research interests lie in the modeling, algorithmic, and systems issues in wireless, vehicular, embedded, and sensor networks. His research has been an integral part of several NSF and DARPA projects such as the GENI WiMAX and the ExScal projects. He is a recipient of the NSF CAREER Award. (URL: http://www.cs.wayne.edu/~hzhang).

Appendix XXVI. Justin Berg, *The IEEE 802.11 Standardization, Its History, Specifications, Implementations, and Future*, Technical Report GMU-TCOM-TR-8.

Graduate Program in Telecommunications
George Mason University
Technical Report Series

4400 University Drive
MS#2B5
Fairfax, VA 22030-4444
http://telecom.gmu.edu/
703-993-3810

The IEEE 802.11 Standardization
Its History, Specifications, Implementations, and Future

Justin Berg
jberg2@gmu.edu

Technical Report GMU-TCOM-TR-8

Abstract

The IEEE 802.11 Standard for Wireless LANs has had a profound impact on the provision of network access and resources to dispersed, and many times varied, network elements. It has not, however, been a static system implemented since its initial ratification. The standard has been under constant amendment and updating, striving to provide new services and capabilities for expanding wireless needs, and address shortcomings in the original standard. From its humble beginnings focusing on interoperability with the broader 802.x standards to provide bridging across various media, to its ongoing search for new RF techniques and spectra for increased support to new applications, IEEE has endeavored to stay ahead of users' requirements for Wireless LAN communications.

1. History of IEEE 802.11

In 1985, the Federal Communications Commission (FCC) deregulated the spectrum from 2.4-2.5 GHz for use by the Industrial, Scientific, and Medical (ISM) communities. This meant that the spectrum would be available for individual, non-licensed applications [1]. This news was exciting to up-and-coming developers of wireless communications technologies, because they could now develop without spending money on licensing fees. Unfortunately, this led to many developments that were far from the ubiquitous, sprawling networks we see now. At the time, and throughout the development of the 802.11 standard, if wireless network technologies were available, they were usually proprietary, expensive, slow, or simply lacked widespread availability/adaptation – and most suffered from several of these challenges [2].

In the early 1990s, however, the IEEE realized that a wireless communications infrastructure standard was necessary to meet a clearly-desirable market niche. The IEEE established an executive committee, as part of the IEEE 802 standard for Local and Metropolitan Area Networks to focus on developing a wireless LAN standard [2]. The 802.11 committee focused on providing a reliable, fast, inexpensive, robust wireless solution that could grow into a standard with widespread acceptance, using the deregulated ISM band from 2.4-2.5 GHz.

The original standard, ultimately adopted in 1997, is vastly different from the standard that exists today. The maximum data rate was 2 Mbps. It included forward error correction, and two forms of interference mitigating spread spectrum methods – direct sequence and frequency hopping. It also included a specification for infrared wireless communications, still operating at up to 2 Mbps.

A large part of 802.11's success is its inherent compatibility with current 802 networks, specifically the 802.3 wired Ethernet networks [2]. The independence of physical access (PHY) and media access (MAC) from overlaying communication layers is critical to this compatibility. This compatibility was part of the 802.11 committee's charter [1], but its implementation played a large role in ongoing internetwork growth. The compatibility was built on two pillars – physical layer compatibility and media access layer compatibility. The separation of these layers is critical to, not only the early implementation of the standard, but the ongoing extensibility of the standard.

The physical layer portions of the original standard, and as well as today's standard, focus on allowing the base stations to get wireless broadcasts to one another; transceiving. The broadcast frequencies were in the 2.4 GHz to 2.483 GHz range or in the infrared spectrum (IR) (850-950 nm) [2]. Transmitters used time-division duplex (TDD) radio broadcasts, allowing both uplink and downlink to share the same RF channel, using differential binary phase shift keying (DBPSK) or differential quadrature phase shift keying (DQPSK) signal modulation (Appendix A). Transmitters used either Direct Sequence Spread Spectrum (DSSS) or Frequency Hopping Spread Spectrum (FHSS) for interference mitigation. Data rates were specified for both 1 Mbps and 2 Mbps operation.

The media access layer (MAC) processes the PHY layer signals into the ubiquitous network layer. Fundamentally, 802.11 uses collision sense media access with collision avoidance (CSMA/CA) for its media access protocol (Appendix B) [2]. The MAC layer also provides several services to assist in the wireless broadcast such as synchronization, power management, frame fragmentation, and frame encryption (WEP - Wired Equivalent Privacy) and authentication, with varying methods of employing these services for both infrastructure-based in distributed (known as ad-hoc) networks. For example, in an infrastructure network, synchronization is performed between all transceivers by using beacons transmitted by the access point. In an ad-hoc network, however, the synchronization responsibility falls to all members of the independent network, creating a sub-network of synchronizers.

Note that there is no 5 GHz spectrum specification in the original 802.11-1997 standard. This frequency allocation was not explored (or at least, published) until shortly after the original standard was adopted. The original standard focused on exploiting the recently-unlicensed 2.4 GHz ISM band, and the practical, and already-in-use infrared spectrum. In fact, the original standard largely overlooks, or at least actively ignores, many compatibility standards that would end up being crucial to widespread acceptance of the standard. For example, the entire standard makes only cursory mention of MAC address space, pointing out that its 48-bit address space is compatible within the broader scope of the IEEE 802 address space, but is not required to be unique from a global 802 address overlay. This compatible address space, which is still a part of the 802.11 standard today, allows 802.11 networks to interact with the 802.1 LAN specification that provides for bridging between separate physical networks, and is perhaps the cornerstone of the success for the standard. This address compatibility with 802.x networks (and flexibility) played a role in the widespread adoption and interoperability of 802.11 wireless networks [2], even in the face of other, higher-speed competing network standards such as HiperLAN, a competing European standard for wireless network communications, which provided its own convergence to internet protocol (IP) networks, vice relying on 802.1 for internetwork bridging [1].

Despite not having addressed direct compatibility of the 802.11 with 802 networks, the committee left the door open, and in fact immediately fostered the follow on Task Groups to address specific supplemental topics for use within the 802.11 standard framework. The 802.11b task group, TGb, addressed higher speed transmissions within the WLAN environment. The 802.11b Task Group produced the 802.11b amendment, adopted by IEEE in 1999, just two years after the original standard was adopted. It allows for 5.5 Mbps and 11 Mbps data rates, using Direct Sequence Spread Spectrum (DSSS) transmissions [2]. It also prompted the creation of the Wireless Ethernet Compatibility Alliance (WECA); a non-profit association for standardization and promotion of Wi-Fi technologies. From wi-fi.org [5]:

> "The Wi-Fi Alliance is a global non-profit industry association of hundreds of leading companies devoted to seamless connectivity. With technology development, market building, and regulatory programs, the Wi-Fi Alliance has enabled widespread adoption of Wi-Fi worldwide."

Even today, 802.11b is probably the most widely-recognized, and widely-used 802.11 standard, although 802.11g is quickly surpassing it, with 802.11n up-and-coming in popularity and availability. WECA renamed itself to the Wi-Fi Alliance in October, 2002 [4].

About the same time the 802.11b Task Group was designing the 802.11b amendment, the 802.11a Task Group, TGa, was doing the same for another wireless standard [3]. At the time, many countries had recently opened up some 5 GHz spectrum for unlicensed (but still regulated) use. This spectrum was less "RF dense" than the 2.4 GHz spectrum [2], which includes other interferors such as garage door openers, cordless telephones, microwave ovens, and baby monitors. With less interference high bandwidth available, another, higher capacity standard could be constructed.

The ultimate 802.11a standard included a 54 Mbps data rate using the more-complex orthogonal frequency division multiplexing (OFDM) waveforms (Appendix C), and operated in the 5 GHz range, set aside for the Unlicensed National Information Infrastructure (U-NII) usage [1]. While the standard was completed and adopted in 1999, the more-complex equipment did not begin shipping until 2001.

It is significant to note that while data rates were increased by both 802.11a and 802.11b, that both only increased data bandwidth within RF applications. The IR specification, while still valid, was left behind with 1-2 Mbps maximum throughput, while the RF environment has continued to increase in data throughput throughout the development of the 802.11 standard.

Not long after 802.11a was adopted, IEEE immediately recognized that the OFDM waveform could benefit the 802.11b standard. Increased data rates would even support bandwidth-hungry multimedia applications as the demand for these applications grew. In July 2000, the 802.11 Task Force G was assigned the task of overlaying the OFDM waveform on the 2.4 GHz spectrum, producing a new standard that was fully backward-compatible with the 802.11b standard. This was no easy feat, but after 3 years the new standard was ratified. The key was in requiring all 802.11g equipment to support complimentary code key (CCK) modulation as a fall-back mechanism to ensure 802.11b compatibility. This fall-back has significant impacts on the total data rate of the network, but allows mixed 802.11b-802.11g network equipment to coexist on the same topology. As 802.11b equipment is phased out and replaced with 802.11g equipment, users can seamlessly upgrade their network without upgrading the entire infrastructure. In June 2003, the amendment was ratified.

As 802.11 enjoyed widespread adoption by home and business users alike, more scrutiny was placed on security. The initial standard included a MAC-level security protocol called WEP, Wired Equivalent Privacy [6]. WEP was intended to provide confidentiality and authentication for connecting users. By using a very small subset (up to four) of pre-shared keys, a user could identify itself as a valid user to an access point, and encrypt every packet of the session [7]. The intent of WEP was not to be a bulletproof security protocol for wireless networks, but to provide reasonable session privacy, like that which could be expected from a direct-connection (wired) connection. Unfortunately, WEP was rife with vulnerabilities (Appendix D), and continued bad press caused 802.11 users to demand better security [7]. Another task group, Task Group I, was set up to address MAC-level security in an effort to address security problems with WEP [6].

The Task Group model, however, took too long to address the concerns of equipment manufacturers. The Wi-Fi Alliance began implementing additional security enhancements to provide customers with additional security features. Many members of the Wi-Fi Alliance were part of Task Group I, and these enhancements would be seen as part of the final 802.11i amendment. These original security implementations, labeled Wireless Protected Access included many enhancements to address the weaknesses of WEP, including the use of extended initialization vectors (IV) (56-bits), rotating initialization vectors, more robust integrity checks, and protection against replay/redirection attacks [6].

In June 2004, the 802.11i amendment was ratified. The security enhancements in it became known as WPA2, Wireless Protected Access v2. It was largely a mirror of the WPA enhancements from the Wi-Fi Alliance, with some small, but significant, improvements. First, it incorporated the use of the Advanced Encryption Standard for encrypting and protecting data [8]. The AES was selected/adopted by the National Institute of Standards and Technology (NIST) in November 2001, and was not available when WEP was being designed nearly 10 years earlier. Next, enhanced integrity checks leveraging the AES CCMP (counter mode with cipher block chaining with message authentication code protocol, a recursive acronym) provides additional authentication. 802.11i also supports several implementations of using external authentication mechanisms, including 802.1X authentications and/or RADIUS [8].

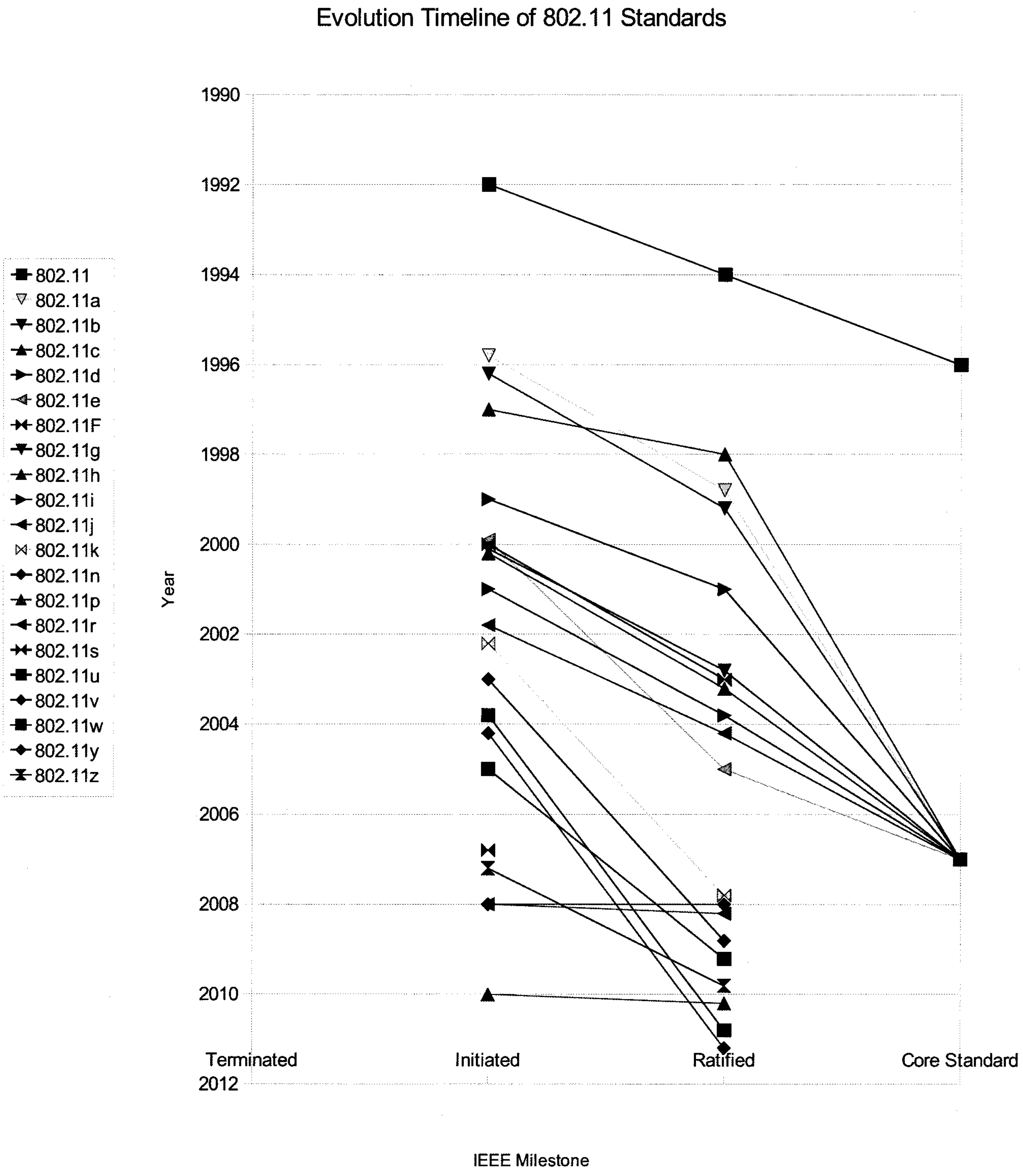
Meanwhile, the IEEE was going through another exercise to increase wireless data rates. Recognizing the seemingly unquenchable bandwidth thirst of users, the IEEE set out to exceed 54 Mbps as an upper data rate limit by creating Task Group n (TGn) in September 2003 [3]. By using multiple-input multiple-output (MIMO) transmitting methods, 802.11n would allow multiple data streams, separated spatially, to increase the overall data rate [9]. This access method, as with 802.11g, is backward-compatible with previous 2.4 GHz implementations of 802.11, as well as 802.11a in the 5 GHz and 3.7 GHz spectra (802.11a was extended to 3.7 GHz by the 802.11y amendment in Nov 2008) [9].

While 2.4 GHz implementations include the largest number of users worldwide, unfortunately the 2.4 GHz spectrum is heavy on interference. While MIMO can provide additional and higher data rates, and protection against some interferences (Appendix E), there is a limit as to how much data can be transferred in the congested spectrum. The 802.11n amendment, ratified in September 2009, can support data rates up to 600 Mbps, but in its current implementation, with the congestion in the 2.4 GHz spectrum, the maximum supported transmission rate is 104 Mbps. This is still a significant increase over the 802.11g amendment, but leaves significant room for growth, should 802.11n be deployed in other RF environments. Indeed, as the amendment does not specify the exact spectrum, the largest performance gains will be realized in the 5 GHz and 3.7 GHz ranges, where significantly less interference is found.

Also, in 2003, the IEEE began consolidating the standard amendments from the Task Groups, and rolling them into a consolidated baseline standard. The standard up to this point had been known as IEEE 802.11-1999. After many amendments, 802.11a, 802.11b, 802.11d, 802.11e, 802.11g, 802.11h, 802.11i, 802.11j, were all rolled into one consolidated standard – IEEE 802.11-2007. While technically these amendments no longer exist, as they are now part of the baseline standard, most still refer to them by their parent amendment designations to easily identify the specific capability or function of the 802.11 wireless LAN standard [9]. The following illustration is a timeline of when specific amendments were initiated (i.e., the Task Group was formed), and when the amendment was ratified, terminated, or rolled into the core standard.

2. The 802.11 Standard

The 802.11 standard, while a single standard, has many manifestations that allow wireless network access. It covers everything from how synchronization should be performed, to how infrared (IR) wireless networks should be configured, to spread spectrum chip rates for different applications. This paper cannot touch on all portions of the standard. Indeed, the 1200+ page standard (not including its many several-hundred page amendments/enhancements) will require this paper leave many topics unexplored, and many, many more topics completely undiscovered.


Evolution Timeline of 802.11 Standards
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
Year
Terminated
Initiated
Ratified
Core Standard
IEEE Milestone
802.11
802.11a
802.11b
802.11c
802.11d
802.11e
802.11F
802.11g
802.11h
802.11i
802.11j
802.11k
802.11n
802.11p
802.11r
802.11s
802.11u
802.11v
802.11w
802.11y
802.11z

2.1 Modes

802.11 networks can operate in two basic modes: infrastructure and ad-hoc[2].

Infrastructure: In infrastructure networks there are two entities: a station (STA) and an access point (AP). This mode is called the Infrastructure Basic Service Set, or just BSS. Access points provide wireless access to network resources for stations, as well as other services such as synchronization and channel selection. Many times, APs are gateways, and provide other services such as network address translation, dynamic host control protocol (DHCP), and other network services. While these services are outside the bounds of 802.11, they are critical to end-to-end usability of the network [2].

Ad-hoc: In ad-hoc networks, there is no access point, just stations. This mode is called the Independent Basic Service Set, or IBSS. Each station communicates and negotiates directly with other stations, with the roles of the Access Point being performed autonomously by the individual stations. This allows a network to be quickly erected and torn down, without the need for in-place infrastructure, thus reducing cost and adding flexibility. While this is not the typical deployment of an 802.11 network, it has its place in the list of 802.11 implementations [2].

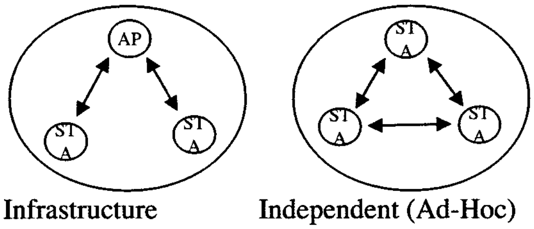


Figure 1: Infrastructure vs. Ad-Hoc (Independent) Modes

2.2 Physical Layer

The biggest changes in the 2.4 GHz wireless network arena has been in the physical layer implementation of the standard. There are many reasons for this. First, the basic access data rate is largely determined by the physical modulation scheme. For data rates to increase, new (more complex) waveforms have been necessary to increase the data rates for the network [9]. Second, since the 2.4 GHz spectrum is unlicensed, there are plenty of interference-producing devices that operate in this range. To avoid, or at least minimize, this interference, spread spectrum methods have been implemented to mitigate these challenges [10]. As data rates and waveforms have changed, new spread spectrum methods have been employed [2].

While the methods used to access the same spectrum may change over time, the unlicensed ISM band has not. IEEE subdivided the entire allocated spectrum into subchannels for use by Wireless LAN transceivers, which are applicable in different regions. Table 1 shows the complete spectrum allocation in the 2.4 GHz range, along with the international regulatory body recognition of the channel allocation [12].

Since the IEEE has continued to provide backward-compatible amendments to the original standard, all the 2.4 GHz methods used by 802.11 for RF transmission are still valid [2]. That is, since no amendment nullifies a previous access method, all the previous methods are still legitimate waveforms and spread spectrum techniques. For a discussion of RF waveforms, see Appendices.

		Regulatory Body (Region)			
Channel Number	Frequenc y (GHz)	FCC	IC (Canada)	ETSI (Europe)	Japan
1	2.412	X	X	X	X
2	2.417	X	X	X	X
3	2.422	X	X	X	X
4	2.427	X	X	X	X
5	2.432	X	X	X	X
6	2.437	X	X	X	X
7	2.442	X	X	X	X
8	2.447	X	X	X	X
9	2.452	X	X	X	X
10	2.457	X	X	X	X
11	2.462	X	X	X	X
12	2.467			X	X
13	2.472			X	X
14	2.484				X

Table 1: 24. GHz Channel Allocations by Region ([reproduction, 12]

802.11 – Allows for frequency hopping spread spectrum (FHSS) and direct sequence spread spectrum (DSSS) to provide interference mitigation and binary phase shift keying (BPSK) or quadrature phase shift keying (QPSK) to provide 1 or 2 Mbps data rates, respectively. Since all future amendments would be backward-compatible, all future methods support these access rates [9].

802.11b – Uses Direct Sequence Spread Spectrum (DSSS) with overlapping channels to provide interference mitigation, and moves away from FHSS for higher data rates. 802.11b also uses Differential QPSK (DQPSK) or Differential BPSK (DBPSK) to provide 11 Mbps or 5.5 Mbps data rates, respectively [2].

802.11g – Uses Orthogonal Frequency Division Multiplexing (OFDM) to mitigate interference and provide 54 Mbps data rates. Different data rates are supported through different modulation schemes, including 64 and 16 QAM (Quadrature Amplitude Modulation). Table 2 shows how various modulation schemes are coupled with transmission types to accomplish specific data rates for 802.11g. The spread spectrum channels overlay with the 802.11b channel layout (see Table 2, below) [14].

Data Rate (Mbps)	Transmission Type	Modulation Scheme
54	OFDM	64 QAM
48	OFDM	64 QAM
36	OFDM	16 QAM
24	OFDM	16 QAM
18	OFDM	QPSK1
12	OFDM	QPSK
11	DSSS	CCK2
9	OFDM	BPSK3
6	OFDM	BPSK
5.5	DSSS	CCK
2	DSSS	QPSK
1	DSSS	BPSK

Tabel 2: 802.11g Transmission Characteristics [reproduction, 14]

802.11n – Uses OFDM to mitigate interference, while using multiple-input, multiple-output (MIMO) transmissions to increase the total data rates. For a primer on MIMO, see Appendix E. MIMO has significant long-term consequences for 802.11. The use of multiple data streams to increase total throughput can dramatically increase the total capacity of the 802.11 environment. Currently, the 802.11b/g channels are 20 MHz, with 1 MHz guards. Table 3 shows the new data rates when MIMO is used with various guard intervals (time intervals between transmission of consecutive symbols) and 20 MHz channel widths [15].

802.11n, however, supports both 20 and 40 MHz channels to provide a maximum data throughput of 300 Mbps. Table 4 shows a significant (doubling) increase in data rates, when 40 MHz channel widths are used. Unfortunately, the 2.4 GHz spectrum is not really arranged in a way to easily take advantage of the higher data rates. In the current channel allocation, there are only 3 discrete channels (non-overlapping) in the American 802.11 standard at 2.4 GHz (channels 1, 6, and 11). In current, energy dense spectra, these new data rates cannot be easily implemented effectively, despite being theoretically possible to implement. [15]

Spatial Streams	Modulation Scheme	Data Rate (Mbps) (GI=800ns)	Data Rate (Mbps) (GI=400ns)
1	BPSK	6.5	7.2
1	QPSK	13	14.4
1	QPSK	19.5	21.7
1	16-QAM	26	28.9
1	16-QAM	39	43.3
1	64-QAM	52	57.8
1	64-QAM	58.5	65
1	64-QAM	65	72.2
2	BPSK	13	14.4
2	QPSK	26	28
2	QPSK	39	42.8
2	16-QAM	52	57.8
2	16-QAM	78	86.7
2	64-QAM	104	115.6
2	64-QAM	117	130
2	64-QAM	130	144.4

Table 3: 802.11n Transmissions with 20 MHz Channel Width [reproduction, 15]

This channel allocation is unfortunate, but all is not lost for 802.11n. First, the standard includes mechanisms to allow Access Points and Stations to identify when 40 MHz channel usage is practical, allowing maximum throughput when feasible, but being able to step down to 20 MHz channel allocation when necessary. In many ways, this is just flow control for frequency division multiplexing. This method protects legacy 802.11b/g networks, as well as other 2.4 GHz transmitters, like Bluetooth [16].

Another method is for the 802.11n station to announce on both 20 MHz channel allocations that all legacy (non-high throughput) network equipment should leave the channel open for some period of time, and then broadcast at the full data rate on both 20 MHz channels (providing 40 MHz channel allocation) for the specified period of time [16].

Additionally, 802.11n is an overlay for 802.11a. Since 802.11a operates in the 5 GHz range, and the spectrum allocation is different, there happen to be 23 discrete 20 MHz channels on which an 802.11n broadcast can capitalize and maximize total throughput. Furthermore, these channels typically have less interference to start with, so channel collisions are less likely to occur [16].

2.3 Media Access

Unlike physical access, media access has not changed all that much for 802.11 since its start. While most 802.x networks use carrier sense multiple access with collision detection (CSMA/CD), 802.11 uses carrier sense multiple access with collision avoidance (CSMA/CA) [2] (Appendix B). This makes some sense as collision detection is easy when there are only two broadcasters on the media. When every station shares the same spectrum for a given access point, however, this is more complex. First, there are more broadcasters, so there are more opportunities to have collisions. Additionally, is the so-called "hidden node problem." [2] In this instance, two stations are broadcasting to the same access point, but due to physical topology, the two cannot receive one another's broadcast – so there is no mechanism to determine (autonomously) that they are having a media collision. See Figure 2 for a graphical representation of the hidden node problem.

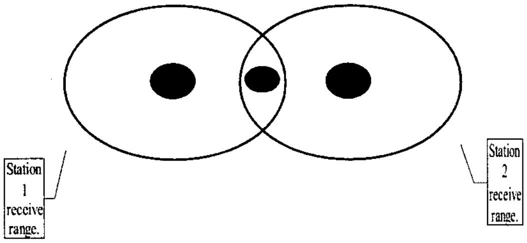


Figure 2: Hidden Node Problem [2]

2.4 IEEE 802.11 Amendments

802.11a – A 5 GHz operating frequency allowing communication up to 54 Mbps using an OFDM waveform [2]. This standard, with its high bandwidth, operating in the reasonably unused 5 GHz frequency band was expected to be the workhorse for business and office applications. While it did find some success in this arena, its high operating frequency was readily absorbed by physical impediments such as walls, floors, and ceilings, significantly decreasing its effective range. Also, the increase in the inexpensive 802.11b hardware made it difficult to justify the the purchase and support of the more expensive 802.11a equipment. Many applications opted for the 802.11b (see below) standard , with lower bandwidth, cheaper hardware costs, and typically further operating ranges.

802.11b – A 2.4 GHz operating frequency (unlicensed ISM), communicating up to 11 Mbps using a Complimentary Code Keying (CCK) broadcast. This standard was expected to be adopted by private individuals and small operating environments due to its inexpensive production costs and relatively convenient configuration options. Unfortunately, the unlicensed ISM band at 2.4 GHz is a "crowded" spectrum, and there are plenty of interference producing products on the market, to include microwaves, portable phones (now, almost completely digital, and nearly defunct), and Bluetooth headsets.

802.11c – Network bridging procedures for compatibility with other 802 networks (specifically 802.1d).

802.11d – Compatibility and conformance extensions for transmitter operation outside of the typical political subdivisions of the 802.11 standard. This effectively is a standard that allows a transmitter to roam between regions that observe/enforce different effective spectra for 802.11 transmitters [17].

802.11e – Quality of Service (QoS) functions for multi-application support. With these extensions, 802.11 is able to support many delay-sensitive applications which might otherwise suffer adverse effects of not having QoS, such as real-time video or voice over IP (VoIP) [3].

802.11f – Inter-Access Point protocol used for managing the handoff of a user between access points in an 802.11 network. This specification was withdrawn in Feb 2006 [3].

802.11g – A 2.4 GHz operating frequency (unlicensed ISM) , communication up to 54 Mbps using a OFMD broadcast, and fully backward-compatible with 802.11b. This specification allows a significant increase in network throughput for the 2.4 GHz range, and when coupled with its backward-compatible feature and inexpensive hardware has lead to a global adoption as the de-facto preferred wireless network standard for "typical" computer applications [14].

802.11h – Extensions to the 5GHz broadcast of 802.11a that address mitigating issues from interference on the 5 GHz spectrum from other applications such as radar and satellite broadcasts, including power control and dynamic frequency selection [18].

802.11i – Security enhancements (WPA2) to address the shortcomings of previous security protocols (WEP). This standard includes support for robust encryption by including the AES, and more robust authentication [6].

802.11j – Broadcast standard to allow 802.11 to be formalized in Japan, where broadcast requirements only allow 802.11 to operate in the 4.5-5 GHz spectrum [19].

802.11k – Radio resource measurement for 802.11 networks, allowing mobile stations to dynamically identify which access points (APs) within range will provide the best network performance, based on more than simply received signal strength, but on a complete report from each AP of its current network utilization, availability and received signal strength [20].

802.11n – The incorporation of multiple input multiple output (MIMO) and increased bandwidth channels to the current broadcast sets (802.11g and 802.11a), increasing the total data throughput [9]. Contrary to current marketing, 802.11n is not a replacement to the 802.11g standard, but is a complementary performance increase that allows the use of multiple (usually two for current commercial availability) 802.11g broadcasts to share the same spectrum, thus increasing (doubling in the case of two data streams) the effective 802.11g data throughput [15]. The same technique can be applied to 802.11a in the 5 GHz spectrum. The current 802.11n standard is designed to support up to 600 Mbps throughput [15].

802.11p – Extensions for vehicular (fast-moving) 802.11 access and handoffs, called Wireless Access for the Vehicular Environment (WAVE). This standard operates in the 5.9 GHz range and is designed to work with a specific overlay infrastructure designed specifically to support such vehicular access needs, such as fast transitions and associations [21].

802.11r – Extensions for fast transition of a user between access points, where the mobile subscriber is in charge of determining the handoff between base stations. While the 802.11 standard addresses handoffs, the increased overhead of authentication and encryption, coupled with increased data rates and latency requirements of real-time multimedia applications (specifically VoIP and to a lesser degree video-teleconferencing), the need for a faster managed handoff is critical to Quality of Service for these applications. 802.11r provides the mobile subscriber (STA) the ability to identify how and when a handoff will occur to maintain performance quality [22].

802.11u – Extensions that allow more seamless movement of a mobile subscriber between 802.11 networks and other external networks, e.g. cellular networks. This is a complex problem, with many facets, from RF compatibility to vertical handoff management to authentication considerations. 802.11u, however, provides a common set of abstraction layers for which compatible external networks can/should be configured to allow interoperability [23].

802.11v – Provides extensions for mobile subscribers to share network topology information, including RF information, so clients can optimize wireless network performance [24].

802.11w – Provides services to protect network management frames [25].

802.11y – Includes standards for operating wireless networks in the 3.6 GHz range, for use in the United States [26].

802.11z – Extensions that allow for direct link setup (DLS) enhancements, creating direct peer-to-peer tunnels,, making the link setup independent of the access point and providing power saving features [27].

3. Security in wireless networks

One of the biggest challenges in wireless networks is security. Indeed, in any network, one of the biggest challenges is security. Security is the bane of both the network administrator, whose job it is to implement security and is held responsible for its failings, as well as the network owner, who is ultimately responsible for the safety and security of the network's data and resources. It is a hard problem, with ever-evolving challenges and nuances. 802.11 has had many iterations of security throughout its lifetime [6]. This section is intended as a security primer to help orient the reader to a security philosophy, and to help demonstrate security as it applies to network infrastructures, specifically 802.11.

3.1 What is Security?

"Security is a process."

While the 802.11 wireless standard contains many hooks to provide security, these hooks cannot be viewed as comprehensive. A wireless broadcast can be completely secure against attack, but if the user hooks up through a wired connection to the network, then everyone on the perfectly-secure wireless connection is at risk if the wired connection is improperly implemented. The 802.11 standard has a series of smart, built-in mechanisms to control security at the RF level (that is, over the wireless network), it neither provides nor prescribes any method to secure the totality of the data network. 802.11i, discussed later in this section, is simply a piece of the comprehensive puzzle necessary to maintain the network security perimeter.

"Security is evolving."

In many ways, this is completely obvious, but it is easy to think about security as an on/off switch, and after finding a "secure" solution to the network risks, then the job is done. Flying in the 1960s was a secure method of transportation. In the 1970s, a slew of hijackings changed the security posture to address clear deficiencies in the airport/aircraft security model. This model met new challenges in the 1980s with the bombing of Pan-Am flight 103 over Scotland, so it was again redesigned. In 2001, a new attack/threat emerged, and again the security model was re-addressed. This is an important concept in security, in that it is important to recognize that security models need to be flexible enough to address changing/dynamic conditions, and need to be fungible to accept new methods to address/defeat evolving threats. That is to say that airport security needs to be able to stay functional during a power outage, and it needs to be able to accept modest changes in procedure and technology to address the latest threats as they emerge.

The 802.11 standard, since its initial

ratification, has faced security challenges, and the committee and indeed the community at large, has constantly striven to address the changing security landscape to keep wireless networks secure. It has addressed changes in the security industry, and left itself a flexible method to continue to adapt as changes become necessary.

"Security is more than just the lock on the door. "

Security can fail in one of two ways – it can fail to keep the bad guys out and it can fail to let the good guys in. It is important to remember the latter, as it is easy to forget that a car that won't let you get in is just as useless as a car that let's anyone get in and drive away.

Furthermore, security can have two failure modes; active and passive. A passive security failure is when the bad guy gets through, despite the fact that the system is supposed to stop the bad guy from getting through. When a car thief breaks the window of the car and the car alarm does NOT go off, this is a passive failure. An active failure is akin to a false positive. When the car alarm goes off because a loud motorcycle rides by, the system is actively responding to a stimulus that is not an actual threat. Sometimes active failures can be far more catastrophic (or at least, lead to catastrophe) than passive failures, because active failures frequently lead to complacency.

802.11 is a standard to provide network ACCESS. It can be implemented in such a locked down implementation that users cannot get access when needed, and is not flexible enough to allow new stations access upon request. This may meet the need for access control, but it probably does so at the cost of network flexibility, i.e. it keeps the bad guys out, but fails to let the good guys in. These tradeoffs (restriction vs. reliable access) will be contrasted in the next section.

"Security should be viewed holistically."

It is relatively easy to focus on one part of security and make the individual segment secure. Unfortunately, security is a weakest-link proposition. Just because a house has a steel door with a high-security lock, does not mean the house is secure. If the house has unlocked/open windows, then a burglar will simply come in through a window, because it is a weaker entry point than the door. Likewise, a house with a solid door, good lock, robust windows, security alarm, and a dog, in most cases does not need the addition of external flood lights if it sits directly beside a house with open windows. A burglar will go for the house with the open windows and forego the hassle of dealing with the security system and alarm (and additional flood lights), and just go for the neighbor's house – unless the secure house is known to contain the Hope Diamond. This is why home security companies provide yard signs to subscribers – it advertises the existence of the home security system. Security needs to be considered in its totality – from assets to equities, from costs to payoffs, and from risks to gains – it is all tradeoffs, and it is all important to the security of the system.

802.11 provides methods to secure the RF broadcast, and even some of the underlying communication (frames and station authentication). These are an excellent start to a comprehensive security strategy, but they cannot be viewed as adequate to secure the network perimeter. Additional mechanisms must be put in place to address other weaknesses in the network.

"Layer security for best effect."

Eventually, a security mechanism will

fail; either passively or actively. Putting layers of security mechanisms in place allow one security mechanism to fail, while allowing another to perform its job correctly. Airport security does not simply rely on the security checkpoint. A failure, active or passive, at the security checkpoint will allow a malicious participant free reign over the airport utilities. No, the addition of access card readers for access to the runways, Air Marshals on flights, cameras in the terminals, and vigilant passengers (and airline employees) are all secondary measures against the first line of physical defense at the security checkpoint. The application of layers of security provides a series of safety nets to prevent catastrophic failure of the security system. It is not that catastrophic failures can no longer occur if security is layered, but that their frequency dramatically decreases as the unintended consequences of one security failure are addressed by the inclusion of a secondary measure.

Wireless network security can be viewed the same way. This is closely related to understanding the holistic approach to security, but subtly different. While an AES-encrypted RF link provides security against eavesdropping, a weak or compromised key can undo this security. The use of an underlying secure protocol, such as transport-layer security (TLS, formerly SSL), provides an additional layer of security within the RF link, should the RF link be compromised. This additional protection could be the different between an eavesdropper reading all network traffic, including bank transaction information (including usernames and password), a catastrophic result, to the eavesdropper reading all Internet browsing history, a likely less-catastrophic result. Layered security is good security practice.

3.2 How is security measured?

In this instance, measured does not typically mean in any kind of numerical or quantitative way. Keeping the above philosophies in mind, there are key goals that security should attempt to address, anytime there is a security discussion, and the ability of security to address these goals is the measure of its effectiveness. It is easy to lose sight of one of these goals when focusing one or both of the others – recall the holistic approach to security in layers.

Confidentiality – Confidentiality deals with ensuring private data remains private, keeping out unauthorized requests to view the data. Working with a doctor or a lawyer, one expects confidentiality of the discussion and any paperwork that is generated.

The use of encryption is another example confidentiality, specifically data confidentiality. 802.11 provides confidentiality through the use of encryption, outlined in the 802.11i amendment, specifically supporting the AES and RC-4 stream cipher.

Integrity – Integrity deals with ensuring, and perhaps validating, data remains unchanged (or at least, marked as changed) in transit or at rest. When someone provides a phone number or credit card number over the phone, frequently the person on the other end reads the numbers back to ensure the data is recorded correctly. This is a very crude form of an integrity check. Notice that it can be used in conjunction with confidentiality, but it is not a requirement. The phone call is not encrypted when a credit card number is provided over the phone, but an integrity check is performed.

Integrity checks are common in many data transactions. By computing an integrity check on every packet, and appending that integrity check within a [defined] part of each packet, the receiver can determine if the packet was disturbed/tampered in transit. This is certainly not unique to 802.11, and in conjunction with forward error correction schemes not found in 802.11 is actually very common in many RF standards dealing with packet-based transmissions. Regardless, integrity checks are part of the 802.11i amendment.

Availability – Availability deals with ensuring resources are available when requested by an authorized user. A claim check ensures that only one person can pick up laundry from the dry cleaner, but if the dry cleaner is closed, then the claim check is worthless, at least for the moment, until the dry cleaner opens again.

Availability is, in many ways, a tricky beast to tackle for data networks. It is often at-odds with confidentiality and integrity. Enhanced authentication and encryption impact who can access the network resources and indeed the network performance; encryption can have significant processing requirements and impact device performance. Yet, without network availability, the entire utility of the network is lost.

3.3 802.11i

With the above concepts regarding security in mind, let's look at the 802.11i standard, and how it provides mechanisms to perform the above functions.

Confidentiality:

802.11i provides three mechanisms which can provide data confidentiality.

Wired Equivalent Privacy (WEP) – In its perpetual effort to remain-backward compatible, 802.11i allows for the continued use of WEP as a confidentiality provider. Interestingly, confidentiality was called "privacy" in the original standard, and the 802.11i amendment makes a global change to this term, amending it to "confidentiality." [8] WEP remains a part of the standard today, but is widely recognized as insecure, and many individual industries have prohibited the use of WEP as a security service for specific functions [6] (for example, ATMs may not use WEP in any transaction regarding user data [29]).

Temporal Key Integrity Protocol (TKIP) – This is effectively an enhanced version of WEP to address some of the more-indicting vulnerabilities of WEP [6]. It still uses RC-4 stream ciphering, but addresses a number of issues including weak initialization vectors for per-packet encryption. Since it relies on RC-4, however, it is still vulnerable to many of the attacks to which WEP is vulnerable. It was included in 802.11i to address WEPs weaknesses, while allowing the current network equipment to be able to implement improved security over WEP – equipment that could not handle the processor-intensive functions of AES [6].

Counter Mode with Cipher Block Chaining Message Authentication Code Protocol (CCMP) – This is a custom implementation of the Advanced Encryption Standard (AES)-CCM implementation with 128-bit keys, providing data confidentiality (per packet encryption), along with authentication and integrity on a per-packet basis.

No confidentiality – Open system – 802.11 does allow for unprotected/no-confidentiality transmissions, so long as both the sender and receiver are configured to allow unprotected transmissions [8].

Authentication:

In the original 802.11 standard, WEP provided station (STA) authentication through the use of shared keys with WEP. If a WEP STA could demonstrate it had knowledge of a suitable key during network association, then the STA was considered an authentic network user [7].

In the 802.11i amendment, this shared-key authentication is still considered valid for WEP implementations, but otherwise 802.11 relies on 802.1X using the Extensible Authentication Protocol (EAP) to provide authentication functions for the network [6]. Keep in mind these are network resource authentication issues, and not packet authentication issues (like the CBC-MAC – cipher-block chain-media authentication code), which provides authentication and integrity checks on individual packets for the 802.x network.

Availability:

802.11 addresses availability through uniformity and predictability, providing a single, global standard for 802.11 products to be interoperable, and ensuring that standard is easily implemented in a way that is not so complex that RF transmissions become unreliable or fickle. By allowing for solid standards, vendor wi-fi certified products can reliably interact to provide consistent, reliable availability to end users [5].

4. 802.11 Applications

Once the wireless LAN framework had been designed and standardized, individual markets had to determine how to use the new technology to benefit specific needs. As more and more wireless LAN products came to market, large manufacturers such as LinkSys, Belkin, and Netgear provided increasingly affordable products to make the introduction of wireless networks into nearly every use case a reality.

The core application of the 802.11 standard is the ability to wirelessly connect and share data across physically disconnected users. 802.11 wireless networks provide a ubiquitous data sharing media, which promotes user mobility, ease of connection, and a network infrastructure that is compatible across operating systems and applications. It also provides for easy setup and teardown of the network infrastructure, allowing networks to be quickly erected to support requirements, and quickly removed when they are no longer needed, without having to abandon an expensive, in-place infrastructure. All applications exploit these core applications, but specific applications need additional support from the 802.11 standard to meet specific niche-requirements of the application.

4.1 Corporate/Office Applications

Corporate offices are arguably the primary focus of initial wireless LAN implementations. Offices have dozens of hardware requirements, from desktops and servers, to printers and scanners, to fax and copiers, to even remote inventory management handhelds and PDAs. Integrating these devices into a seamless internetworking infrastructure is a challenge. Many of these devices lend themselves to hardwire connections (like desktops), while others are almost crippled by a hardwire connection (like PDAs), with others still in the middle (like laptops). Wireless networks help bridge the gap between these device differences within the workplace, bringing together one single operating network, which is a clear goal for the 802.11 standard.

A corporate office likely has physical control over its office space in a way that allows for relatively easy wiring (and pre-wiring) for Cat-5 cabling, supporting Ethernet for office applications. Built-to-suit office space, modular construction, and inexpensive materials allow companies to lay out a floorplan, with ample wired data hookups, and to have data ports available to provide for present and future data capacity requirements. Servers, routers, desktops, and printers can all be easily hooked up to the network via hardwire, as the company pre-wires the office space for a semi-static layout (offices do not rearrange their layouts very often), with room for future growth.

This pre-wiring lends itself to static and semi-static arrangements, but does not immediately support forward-leaning portabilty and mobility requirements. Sales and marketing need a portable office, contained on a laptop. Vice presidents need access to PDAs while on the road. Logistics needs handheld scanners to inventory management. Wireless handsets for the office's new VoIP service is a logical extension of a wireless capability. These applications do not fit neatly into a hardwired infrastructure, although some could, admittedly, be smartly and reasonably overlaid with a smart wiring plan (for example, dedicated Cat-5 hookups for every laptop). A wireless LAN is needed to support real-time updating and continued productivity on these devices.

The 802.11a standard was intended to solve this problem, having large bandwidth (54 Mbps, per the specification), but its limited range, higher costs, delayed time-to-market relative to 802.11b equipment and limited vendor competition made the 802.11a capability a little too ahead of its time. As companies started to adopt wireless networks as a standard part of operating, the cheaper, readily-available 802.11b network equipment was reasonably-priced, ready to install, and its data throughput met the meager needs of the market at the time (11 Mbps by specification). And so it was that 802.11b ultimately dominated the business market for wireless, intra-office and intra-plant wireless data communications.

Since the standardization of the 802.11g amendment, with its backward-compatible specification, offices have continued to feed their need for data throughput with follow-on equipment in the 2.4 GHz ISM band, getting 54 Mbps data throughput (by specification). This will surely be the case with the follow-on 802.11n standard, which is currently in draft, but will continue to be backward-compatible with 802.11b/g, and will continue to operate in the unlicensed 2.4 GHz ISM band.

Offices face unique security challenges, also, when it comes to 802.11 networks. Specifically, the need to maintain the integrity of data access on the network. How does a corporation provide the flexibility for users on mobile devices (to include laptops) the ability to access the corporate network, and corporate resources, without allowing a rogue wireless device onto the network. This rogue device could simply borrow the Internet connection (low risk) to steal corporate secrets (med/high risk) to attempt to corrupt, collapse the corporate network (high risk). These challenges, while non-trivial, are surmountable. The layered approach to security (discussed more in the previous security section) can certainly help mitigate these risks, and minimize the chances of a total security failure. Access controls on file systems, Intrusion Detection Systems (IDS), MAC address filtering on users connecting to the network, and user-level authentication can combat a rogue device accessing the network, but what can the 802.11 standard provide?

The 802.11i amendment provides

security features such as encryption and authentication for wireless users through the use of WPA2. These security specifications can be used to secure the network against a rogue device connecting to the network by using either MAC (media access control) address filtering or enhanced MAC (media authentication code) packet-based authentication. Additionally, the use of AES encryption can also be used to protect against a rogue listening device passively collecting and exploiting the frames in transit. The 802.11 security methods have been overhauled a couple times in an effort to update/protect against new attacks against wireless networks, with the flexibility for more updates should they become necessary. While WPA2 does not address all the possible security issues in a wireless network, it addresses the lowest level attack of a rogue station attempting to subvert the network integrity.

4.2 Home Applications

Home applications were an obvious extension of the original 802.11 standard, but perhaps a bit premature in some ways, and probably not a primary focus of the standard. At the time of the 802.11 development, more and more consumers were getting personal computers, but certainly multiple computers per home was not the norm – much like automobiles in 1950s America (cars were becoming standard in a home; multiple cars were a luxury), so there was little need for a "network" with only one terminal. Additionally, Internet access in the mid-1990s was still in its infancy for the home user. Dial-up access was available, with any broadband access only being available to intrepid technophiles who contacted the phone company and leased a larger data connection (T-1) for home use. Indeed, there was little need for networking, let alone wireless networking in most homes.

With the introduction of broadband services in the late 1990s, the explosion of the utility of email and the Internet (remember the dancing baby?), the increasing affordability of desktop computers, and the increase of corporate America's use of laptop computers to perform work from home after hours, a need to internetwork computers in the home, and share broadband internet connections became critical. Wireless networking had its niche opened in home applications.

While the home application is similar in many ways to an office application, there are many differences. While a corporation may have the resources to pre-wire large portions of its office space to allow for simple addition and removal of network devices, a home user probably doesn't have the foresight, know-how, money, time, or desire to pre-wire their home for network connectivity. While some new homes, particularly in the early- to mid-2000s came pre-wired with Cat-5 wiring for Ethernet, the vast majority of homes don't come with this infrastructure. Retro-fitting a comprehensive wiring plan over an older home or rented apartment, particularly one with plaster walls, becomes a remarkably tedious endeavor – one that will challenge the patience of the most handy of handy-men.

This inability to retro-fit a home with wiring makes the wireless network most appealing as the ease of setup, and the ability of a "typical" wireless router (for the cost of ~$50), will easily provide adequate RF coverage of the "typical" American home (~2,000 sf). For a reasonable price, a home user can provide network coverage to anywhere in their home, versus having to select a handful of locations and painstakingly pulling cable to support a wired network. The advantage clearly goes to the wireless infrastructure.

Note that the advantage of mobility is not the primary consideration for most home users. While a laptop may very well be a network device for a home user, it is the desire to set up the network quickly and easily that makes the wireless network appealing to the home user. So, what are the security concerns?

The home user probably is not concerned about security the same way that a corporation is. Admittedly, this is probably out of ignorance in most cases, but also, the value-proposition is not structured quite the same way. The threats are the same – someone stealing/borrowing the internet connection, stealing data, or trying to maliciously destroy the network, but the risks are actually different. A user can only have his/her data stolen when his/her computer is on, as opposed to a corporation that has its server on 24-hours, or at least on 10-hours a day during business hours. Once again, 802.11 authentication and encryption can help protect data in transit, and can keep a rogue device from gaining access to the network or the network data, but encryption is enabled on surprisingly few home networks. A home probably does not have an "IT support" person or team close by to help with network connectivity issues, so a collapsed/failed network has a bigger impact on usability than a corporation.

Certainly, these are broad assumptions about the user class (the IT support people probably live SOMEWHERE, so clearly their homes have an IT support network), but the fundamental idea that the "average" home user is interested in reliability and ease of use over security and flexibility is key to understanding the core security challenge to the average home user – availability.

The biggest security feature the 802.11 standard has to offer the home user is availability – the ability to simply and reliably connect to the network each and every time the user desires. A home user can easily misconfigure their network or laptop to lose network connectivity, and may not know how to solve the problem in the event of a misconfiguration. It may be, at least in the short run, in their best interest (and certainly easier) to prevent the loss of data access by simply not configuring anything that might limit their ability t o access the network [30].

Configuring a wireless internet access point with heavy encryptions and strong authentication mechanisms can mitigate most risks of data loss and unauthorized access, but again, a layered approach to security with a holistic outlook helps address ALL security risks. Data loss and unauthorized access can be mitigated by other mechanisms. For example, data loss can be prevented by additional controls such as static encryption or simply shutting down the desktop whenever it is not in use, and unauthorized access can be significantly decreased by good desktop firewall software (and perhaps even an IDS). So including data protection methodologies in conjunction with an open access point, helps a home user meet all their security requirements, even if they do not realize it. Many home users have antivirus and firewall software "because it's a good idea," and many do not have any protection mechanisms in place on their wireless access point. While they may not have arrived at this solution by taking a holistic approach to security, they may have arrived at a suitably-comprehensive security solution simply by following the industry's common wisdom.

4.3 Hotspot Applications

Hotspot applications were probably always a vision of the original 802.11 authors, but not in their most current implementation. Hotspots are available in so many places, but not available "everywhere" in the ubiquitous way that the authors probably envisioned. I imagine the original authors originally envisioned large portions of downtown areas covered by a ubiquitous single network, getting coverage from dozens of wireless access points. While this may exist in some small urban microcosms, as part of some urban renewal project or a technology outreach program, this is not how wireless hotspots have mostly developed. Many corporate and college campuses have achieved this arrangement, but their application still mirrors the corporate application above, with a known user class and understood devices, more than an open infrastructure for the convenience of a largely unknown population.

Most hotspots today, those which provide wireless network access openly to an undefined population are limited to single establishments providing wi-fi service to patrons for the duration of their stay at the establishment. Coffee houses, McDonald's, airports and airport lounges, hotels, and libraries seem to top the list for hotspot providers. A patron comes into the establishment, establishes themselves as a legitimate customer, and is then provided access to the network resources upon demand.

Unlike home and corporate applications, providing wired infrastructure is not only inconvenient, in many ways it is self-defeating. By definition, the user population is mobile, and the establishment is catering, and in fact promoting this mobility – a coffee shop does not WANT a user to stay for 6-8 hours after ordering a single latte. PDAs, iPhones, Androids, and iPads are an expected network devices for these users, and a wired connection simply will not provide the desired network connectivity. The mobility or the user, and subsequently the wireless network access, is crucial to the business case of the hotspot in most cases.

So while internetworking is a significant driver for corporate applications, and ease of installation is important for home applications, hotspot applications focus on the transient nature of the expected user class. In this environment security becomes a very peculiar being. In most hotspots, because the proprietor does not have an interest on the network (in most cases), i.e. they are not users of the hotspot, data loss is not a critical factor. Additionally, having the internet connection stolen is not immediately a factor, as the internet connection is provided for the very purpose of allowing transient mobile users to access it. So what are the security concerns?

The primary security issues with a hotspot internet connection are accessing the hotspot without being a legitimate patron of the providing establishment (authentication – low risk) or legal issues with the connection of the mobile user (hacking other websites/users, or performing illegal functions, like sharing large volumes of illicitly-obtained music) (med/high risk). The latter is usually protected against from a legal perspective, by requiring users to acknowledge the legal uses of the hotspot – a function that resides well outside the 802.11 purview.

The former, however, can be a significant service issue for a hotspot proprietor. For example, a coffee shop in an apartment building needs to address how to keep apartment-dwelllers from connecting to their network and using all their bandwidth, effectively denying intended service from coffee-drinking patrons who intend on using the hotspot service. Unfortunately, the

802.11 standard does not have anything to immediately address this requirement, and such proprietors need simply find third-party services to help address these issues.

5. The Future of 802.11

So what lies ahead for the 802.11 standard? In many ways, the 802.11 standard is already preparing for the future. Amendments have been included to allow for more formalized decentralization of the network infrastructure [35]. Other amendments are working toward promoting vehicular access [33] , and still others are identifying newer, faster access speeds [37, 38]. They are complex problems with complex solutions, but the future of 802.11 seems to be in good hands. Here are a few of the amendments that will support 802.11's growth into the future of wireless networking.

5.1 Ratified Amendments

802.11p – Support for Wireless Access in Vehicular Environments (WAVE)

802.11 is primarily focused on allowing independent stations to identify and associate with one another, and while 802.11 is built for "mobility," it does not embrace a truly "mobile" user, i.e. a user that is presently in transit. 802.11 can be shoehorned into mobile user and mobile sensor environments, but often dynamic association is difficult, and power and antenna requirements may not be appropriate for many sensor requirements. For the most part, 802.11 does not address how to perform association and data sharing in mobile environments, particularly in high-speed environments.

802.11p provides a physical and media access method in the currently-unused (by 802.11, anyway) 0.8 GHz range, known as the Dedicated Short Range Communications (DSRC) band. While the amendment has hit its stride, yet, being only recently ratified (July 2010), it will support Wireless Access in Vehicular Environments (WAVE), and supports high-speed STA association and data transmission/sharing [33].

This information-sharing will allow vehicles to share information with other vehicles to support a variety of inter-vehicular applications [33], such as collision detection and avoidance, or emergency vehicle detection and alert. One application might be a driver being made aware that a firetruck was on an intercept course with the driver's course within the next mile, instead of hearing it at the next intersection. Another application could be two cars adjusting their distance because one driver was drifting into the lane of the other car. Other stand-alone applications (single vehicle) could also support toll collection at speed, speed monitoring, traffic direction and monitoring, and curfew compliance.

802.11u – Internetwork Support to External Networks

Communications providers have long tried to leverage the overlying digital cellular network with the interstitial wireless LANs. Some cell phone manufacturers have built cell phones that allow a caller to roam from the cellular network to a wi-fi network, and seamlessly transfer to the wireless LAN (IP) network. At least, that was the concept. Time and again these transitions have demonstrated to be dubious, at best. Problems with inter-network authentication, vertical data handoffs, and data services compatibility made what seems like a trivial problem ("bits are bits") into significant integration and handoff challenges.

The 802.11u amendment has

attempted to formalize the process by which providers and equipment manufacturers build networks and network equipment to be interoperable between wireless LANs and overlying RF networks, such as GSM, CDMA, or WiMAX (and IEEE 802.16) [36]. By formalizing the process of network advertisement and discovery, user equipment (STAs), can identify the local network infrastructures. 802.11u allows a STA to tunnel through the network to the appropriate subscriber network to authenticate with the subscriber network, independent of the 802.11 disposition [23]. This prevents the passing of potentially sensitive credentials to the 802.11 network, and alleviates the 802.11 network from having to be compatible with each and every subscriber network with which a STA might want to authenticate/access [36]. 802.11u finally specifies a QoS Map service which allows the subscriber network to identify how a vertical handoff (layer-3, end-to-end services) should be performed to promote seamless intern-network handoffs 23].

While not in widespread use, yet, having 802.11u ratified is a significant step in the process of promoting inter-network communications, which has historically been very difficult. The compatibility of these networks will help address power and battery issues in mobile devices, using low power in the vacinity of wireless LANs (APs) while still receiving significant data throughput, and having access to 4G and beyond celluar data networks with broad geographic coverage.

5.2 Future/Draft Amendments

802.11s – Support for Wireless LAN Mesh Networking

While today the primary implementations of 802.11 are through access point (infrastructure) implementations, the standard does allow for independent stations to communicate with one another in ad-hoc mode. In this mode, both stations effectively negotiate the connection between themselves, creating their own small, direct, private network. Nodes of the network can share information directly, and if one has an internet connection, with proper routing configuration, the internet connection can be shared with other members of the ad-hoc network. Unfortunately, ad-hoc networks suffer from the "hidden node" problem, so it is not guaranteed that a single station can send to every other station that is a part of the ad-hoc network.

The 802.11s amendment, which is still in draft, will allow stations to create full-mesh networks throughout the ad-hoc network. This full-mesh arrangement will allow member nodes, not only to communicate with "over the horizon" nodes, those nodes that are "hidden nodes," but also to transit networks through mesh routing, to reach internet gateways to provide access to the internet, and to simply reach network resources that are not readily available on the ad-hoc network [35].

The 802.11s, as drafted, relies on the Hybrid Wireless Mesh Protocol (HWMP), which is based on the Ad-hoc, On-Demand Vector (AODV) routing protocol. In completely ad-hoc networks, using AODV individual nodes learn about the network topology via a series of route requests and route replies through a tree structure. AODV supports unicast and multicast addresses, and nodes update their routing tables periodically to remove stale routes. An individual node broadcasts a route request every time a destination is required that is not currently in the routing table. Adjacent nodes forward the route request, noting the request has been made by the originator. When a response is returned, nodes along the return path identify the destination node/address/gateway. Sequence numbers are used to identify the "freshness" of a route, and to prevent routing loops. This method allows the system to be scalable, flexible, and self-healing. This method also allows each node to know little about the entire network topology, until there is a need to communicate with foreign nodes, minimizing unnecessary routing updates [34].

In networks with some infrastructure, like an internet gateway, the HWMP can use tree-like routing to create distance vector routes to the gateway to facilitate consistent, efficient routes to destinations outside the mesh network. This hybrid approach (AODV for inter-network and tree routing for extra-network) allows the mesh network to learn about its own topology changes, without having to exhaustively search the internet for routes to public (extra-network) address space [35].

The introduction of 802.11s will have several impacts on potential future applications. Logistics tracking hardware (for example, hand scanners) can share information with a central database, without the need to have a locally-installed infrastructure. Emergency responders can quickly deploy smart networks to support disaster-recovery operations. Soldiers on the battlefield can generate their own, self-healing ad-hoc networks to share information. Effectively, any application where an infrastructure is immobile, impractical, or otherwise unavailable, can be addresses strictly through deployment of multiple 802.11s -compliant pieces of hardware.

802.11ac/ad – Spectrum Allocation and Techniques to Support Very High Throughput

802.11 currently allows for operation up to 450 Mbps with the 802.11n amendment, when pushed to its maximum potential. This increased bandwidth will be a boon for multimedia applications. The 802.11ac and 802.11ad Task Groups are shooting for more, though – much more. Looking for at least 6 Gbps, the Task Groups are trying to identify how HD video can be transmitted to STAs.

Task Group ac (TGac) is attempting to address this high-bandwidth requirement, called Very High Throughput (VHT), by using the existing 2.4 GHz and 5 GHz bands (technically, they're chartered with "below 6 GHz" operation) [3]. Early signs show that the amendment will support (or attempt to support) multi-user MIMO (MU-MIMO) to accomplish the increased throughput in effectively the same spectrum allocation, and coupled with 256-Quadrature Amplitude Modulation (256-QAM), and up to eight independent data streams, reach data rats over 6.9 Gbps [38]. Cell phone manufacturers have already embraced some portions of what may be included in a new standard by including draft-amendment hardware in new phones. The chips are backward-compatible and cell phone manufacturers do not want to end up behind the technology curve [37].

Task Group ad (TGad) is attempting to provide these higher speeds by using a 60 GHz spectrum. The additional bandwidth available in this spectrum makes the 6 Gbps target much easier, in many ways, and precludes the need for channelbonding – something that's necessary for the TGac to accomplish its throughput goals. Channels for TGad may exceed 2 GHz each, compared to the 80 or 160 MHz for TGac. The cost is that the 60 GHz spectrum has remarkably limited transmitting ranges, and may only operate inside of a single room. To help address this range weakness, 802.11ad may ultimately include beamforming standards to increase effective ranges [38]. Only time will tell, but ultimately, users will have significantly more throughput available to them.

802.11af – Reappropriation of TV White Space to Support Wireless LANs.
As TV broadcasts have become more efficient with the requirement to be digital, the increased headroom from channels 2 to 51 leave a significant amount of RF spectrum available to perform more functions. Task Group af (TGaf) has been tasked with defining how 802.11 can benefit from this "available" spectrum, providing increased data rates than current 802.11n transmissions. This spectrum also has the benefit of increased coverage (range) since these VHF/UHF bands propagate further than the current 2.4 GHz and 5 GHz bands [39].

802.11ah – Sub-1 GHz Support to Wireless LANs.
Today the 802.11 standard supports operation in 2.4 GHz, 5GHz, 5.8 GHz, and recently-added, 3.7 GHz spectra. The 802.11ah draft amendment is intended to add sub-1 GHz operation to this list. Different regulatory bodies throughout the World have different available frequencies that are being reallocated and made available for future use. Task Group ah (TGah) has been tasked with identifying how a standard can be constructed to allow wireless LANs to capitalize on the various open environments below 1 GHz in light of the ever-changing, and over-congested sub-1 GHz regulation. TGah's primary focus will be on providing a standard that supports Smart Utility Networks including sensor networks, smart grid (power grid) or remote utility measurement and management networks [40]. It appears that 802.11ah is a long way off, but the idea that a sunny day could cause a house to close its blinds may be a little bit closer after 802.11ah arrives.

6. Conclusions

The 802.11 standard has come a long way from its initial inception in the early 1990s. It started as a method to allow disconnected users to attach to 802.x networks wirelessly, capitalizing on the unlicensed 2.4 GHz and 5 GHz bands. As its popularity increased, however, the need for more bandwidth became apparent. 802.11 still has not been able to satiate users' desire for increased throughput, but the 802.11n standard should at least quell the immediate needs of network applications, providing 100 Mbps service to wireless users.

The bandwidth fix is only temporary, however, as new, bigger, faster, hungrier applications will need even more bandwidth. IEEE is continuing to expand the search for bandwidth from the traditional 2.4 GHz and 5 GHz bands to the 60 GHz band, and looking for over 6 GBps.

While growing as fast as practical, IEEE has kept an eye on security and user privacy; identifying flexible methods to support layered security approaches. Even the Wireless Access in Vehicular Environments amendment (802.11r) allows for privacy concerns to be considered. Truly, though, the 802.11i amendment has been the workhorse for security and confidentiality within the 802.11 standard, and still provides today's security within the wireless networks. When coupled with 802.1X, a layered security approach within the network, 802.11 is afforded increased security robustness.

If there is one thing that has been demonstrated, however, is that the standard is flexible. After demonstrating higher throughput was possible from the humble 1-2

Mbps data rates, there was an immediate move to find better waveforms and hardware. When WEP was demonstrated to be insufficient to protect user privacy concerns, the Wi-Fi Alliance patched the problem until TGi was able to complete an amendment to formally address the security concerns. From operating at data rates of 1-2 Mbps in just a couple frequency bands (IR, 2.4 GHz, and 5 GHz), to operating in sub-1 GHz, 3 GHz, and 5.8 GHz, and beyond, 802.11 in nearly every country, across diverse regulatory domains, the IEEE 802.11 standard has shown it can meet, not only the needs of today's challenges, but it is flexible enough to address challenges in the future.

Appendix A – PSK – Phase-Shift Keying Techniques

Phase-shift keying (PSK) is a data transmission method that allows a carrier signal to indicate digital data by virtue of its transmitted phase. In the case of binary phase-shift keying (BPSK), a digital 0 may be represented by a carrier signal transmitted with no phase change (0 degree, or 0 radians), while a digital 1 may be represented by a complete phase shift (180 degree shift, or π radians – see Figure 3). A receiver simply needs to receive the synchronized carrier signal to determine the phase shift, and then determine the corresponding data value (0 or 1). In the case of BPSK, two data values (1 baud) can be sent per signal (per symbol), since the transmitted signal has two possible phase-shifted states (0 and π radians).

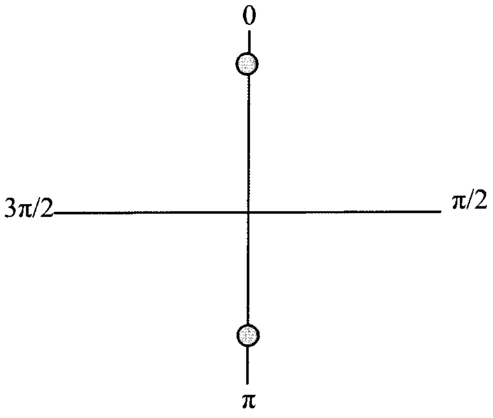


Figure 3: BPSK Phase Diagram

Quadrature phase-shift keying (QPSK) uses the same principle as BPSK, but incorporates the use of two data values per symbol by transmitting a four possible phase-shifts. By transmitting in one of the four quadrants (0, π/2, π, 3 π/4 radians – see Figure 4), two pieces of information can be transmitted per symbol, e.g. digital values 00, 01, 11, 10. This coupling of data allows for faster data transmissions speeds, but also complicates the transmissions. Significant inter-symbol interference(ISI) and fading, for example, have a much larger impact on a simple QPSK transmission over a simple BPSK transmission.

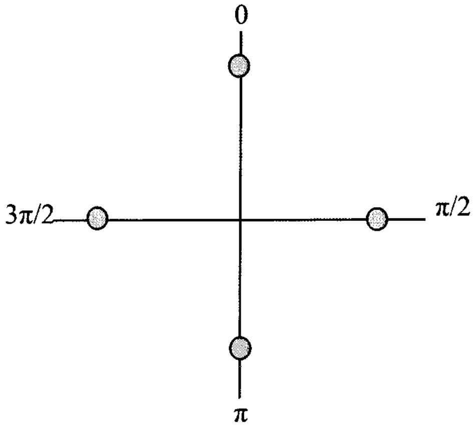


Figure 4: QPSK Phase Diagram

Differential phase-shift keying (DBPSK and DQPSK) is effectively the same method of broadcast, but instead of needing a phase reference for the transmitter and receiver, the phase reference becomes the carrier broadcast. The receiver is able to decode the data values (again, one bit for BPSK and two for QPSK), based on the differential phase-shifting of the received signal. This is more expensive to implement, but solves some issues with fading in RF implementations.

Appendix B – CSMA/CA – Carrier Sense Multiple Access with Collision Avoidance

Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) is a media access method used to allow multiple transmitters to share a single media. In the case of 802.11, the media is the RF spectrum. CSMA/CA allows individual stations to determine if the RF channel is free by having the station listen on the channel to see if it is free.

CSMA/CA process:

1. Determine if the channel is free for a random period of time. (Listen for RF energy in the channel.)
2. Emit a signal to indicate to other stations that the channel is about to be used. (jam signal)
3. Transmit a frame.
4. Listen for another random period of time.

Likewise, other stations are performing the same operation at the same time. In the event the channel is not free during the random wait period, (i.e. another transmitter sends a jam signal and frame during the random listening period) the station simply waits for the other transmitter to transmit its frame by listening to the channel. Once the transmission is complete, the station listens for another random period of time, and tries to transmit when the random window is complete – unless of course another station sends a jam signal.

This media access method allows all transmitters, no matter how many (although there is a practical limit based on the frame size, RF spectrum, and other factors,) to access the same media to perform data transmissions, while minimizing the number of re-transmissions necessary over the media.

Appendix C – OFDM – Orthogonal Frequency-Division Multiplexing

Orthogonal Frequency Division Multiplexing (OFDM)is very similar in many ways to frequency division multiplexing (FDM). In FDM, a single band of spectrum is broken down into several sub-bands. A broadcaster, through whatever mechanism, selects a sub-band, and modulates its data into that sub-band to broadcast.

In OFDM, the same principle applies. A spectrum is broken down into sub-bands, but in the case of OFDM, each sub-band is constructed in such a way that each sub-band is orthogonal to every other sub-band, such that the co-channel interference across sub-bands is effectively zero. Finally, in OFDM, instead of selecting a single sub-band to broadcast on, the transmitter uses multiple [orthogonal] sub-channels to transmit.

To maximize the effectiveness of this method, each sub-band has a complimentary data set transmitted on it, frequently a low data rate transmission. This further helps the signal-to-noise ratio or each channel.

The idea of having multiple data channels to transmit on is not novel or unique, actually. The idea has been around for decades, but there has been no effective way to break data streams up into independent data channels for modulation onto the sub-channel. It has only been in the last decade or so that FFT transmitters have become reasonably-priced to allow OFDM transmitters (and receivers) to be practical.

OFDM can be combined with MIMO (Appendix E) to provide significant spectral efficiencies and data throughput gains.

Appendix D – WEP – Wired Equivalent Privacy

As part of the original 802.11 standard, the Wired Equivalent Privacy protocol (WEP) was supposed to provide privacy to wireless users commensurate with the privacy one could expect from a direct wired 802.x network connection. Wireless connections require broadcasting information to both the receiver AP, but also any other receivers within range, which complicates this level of privacy, since a wired connection is a point-to-point connection (terminal to terminal), unless a bus system is in place, which was rare, even in the late 1990s, while.

WEP uses pre-shared , 40-bit keys between the AP and station first to authenticate the station as a valid user of the AP resources. A station that presents a valid key is considered to be a valid user of the AP resources. Next, the pre-shared key is used in conjunction with the RC-4 stream cipher to encrypt individual bytes of user data before they are transmitted to the AP. Finally, integrity is maintained using a cyclic redundancy check (CRC).

WEP was quickly identified as not being adequate for wireless privacy applications after the 802.11 standard was originally released. 40-bit keys were used to allow international export of the technique, and woefully weak – at the time the U.S. Government had export restrictions on data protection mechanisms. The use of a stream cipher for protection was a methodical mismatch with the wireless broadcasting of 802.11, as any device in range of the transmissions can eavesdrop, providing the eavesdropper with unfettered (and undetectable) access to effectively unlimited datasets about the encrypted transmissions. As one of the strengths of the stream cipher is to have unique codes (key plus an initialization vector (IV)) for every packet, providing unlimited, undetectable access to a third party increases the chance the third party may encounter an IV collision where the same code was used to protect user data. With only four pre-shared keys available, WEP was further weakened by having several stations broadcasting using the same keys, exacerbating the problem of eavesdropping on IV collisions.

The weaknesses of WEP were quickly addressed by the Wi-Fi Alliance with the Wireless Protected Access (WPA) method, which was not part of the standard but simply an industry response to user needs. WPA was finally superseded by WPA2 in the 802.11i amendment, which standardized WPA and incorporated stronger access/privacy methods like AES.